

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 19:06:11 ; Search time 380 Seconds
(without alignments)

268.307 Million cell updates/sec

Title: US-09-831-272-7

Perfect score: 24

Sequence: 1 cagccaccaaaaggagccagaat 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409941 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_29Jan04:/*
2: geneseqn1990s:/*
3: geneseqn2000s:/*
4: geneseqn2001as:/*
5: geneseqn2001bs:/*
6: geneseqn2002as:/*
7: geneseqn2003as:/*
8: geneseqn2003bs:/*
9: geneseqn2003cs:/*
10: geneseqn2004s:/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

ALIGNMENTS

RESULT	ID	Description
1	AAA27970	AAA27970 standard; DNA; 24 BP.
2	XX	
3	AC	
4	XX	
5	DT	15-AUG-2000 (first entry)
6	DE	Box S elicitor-responsive cis-element nucleotide sequence.
7	XX	
8	KW	Box S; elicitor-responsive cis-element; parsley; PR1 promoter;
9	KW	chimeric promoter; pathogen infection; transgenic plant; resistance; herbicide; local response; genetic engineering; disease resistant crop; eli 7; ss.
10	XX	
OS	OS	Petroselinum crispum.
XX	PN	
XX	W020029592-A2.	
XX	PD	25-MAY-2000.
XX	PF	12-NOV-1999; 99WO-EP008710.
XX	PR	12-NOV-1998; 99EP-00121160.
PR	27-AUG-1999;	99EP-00116381.
XX	PA	(PLAC) MAX PLANCK GES FORDERUNG.
XX	PI	Kirsch C, Logemann E, Hahlbrock K, Rushton P, Somsich I;
XX	DR	WPI; 2000-387804/33.
XX	PT	Chimeric promoters mediating gene expression in plants upon pathogen infection, useful for transgenic plant production comprises at least one cis-acting element to direct elicitor-specific expression.
XX	PS	Claim 1; Page 29; 73pp; English.
CC	CC	This sequence represents Box S; an elicitor-responsive cis-element necessary for the expression of the Parsley eli 7 genes. The present invention relates to chimeric promoters capable of mediating local gene expression in plants upon pathogen infection. The chimeric promoters comprise at least one cis element (see AAA27964-A27979) capable of directing elicitor-specific expression, and a minimal promoter. The chimeric promoters are useful for producing a transgenic plant which has attained resistance or improved resistance against a pathogen. The cis-
CC	CC	Ab105460 Drosophil Add4763 Human gen Ab10713 Drosophil Ab10712 Drosophil Continuation (5 of Acc06211 Human cDN Acc06210 Human CDN Acc06213 Human cDN Continuation (17 o Acc31733 Prokaryot Acc10196 Human rep Ab19655 Human rep Ab18726 DNA encod Ab118193 Drosophil Ab118192 Drosophil Ada98924 Human sec Ada45528 Human sec Adc20962 Human sec Ab118192 Drosophil Ada02117 Mouse Nfa Adb7455 Mouse Nfa Ad28763 Human AKA

CC acting element, chimeric promoter, recombinant gene encoding the chimeric
 CC promoter vector comprising the chimeric promoter and a compound capable
 CC of activating the chimeric promoter are useful for producing pathogen
 CC resistant plants, and for identifying and/or producing compounds capable
 CC of conferring induced resistance to a pathogen in a plant. A compound
 CC which specifically activates or inhibits genes activated in a plant when
 CC attacked by a pathogen is also useful as a plant protective agent or a
 herbicide. The chimeric promoter provides rapid and local response to
 CC pathogen attack but shows negligible activity in uninfected parts of the
 CC plants and therefore is most suitable for the engineering of disease
 CC resistant crops

SQ Sequence 24 BP; 10 A; 8 C; 5 G; 1 T; 0 U; 0 Other;

Query Match 100.0%; Score 24; DB 3; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.79; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGGCCACCAAGAGGACCCAGAT 24
 Db 1 CAGGCCACCAAGAGGACCCAGAT 24

RESULT 2

ABX62031/C
 ID ABX62031 standard; DNA; 431 BP.

XX ABX62031;
 AC
 DT 25-FEB-2003 (first entry)

XX Arabidopsis thaliana expressed sequence related polynucleotide #146.

XX Transgenic plant; plant; genetically modified cell; environmental stress; fungicide screening; insecticide screening; gene; ds.

XX fungicide screening; insecticide screening; gene; ds.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX US2002040490-A1.

PN 04-APR-2002.

PP 26-JAN-2001; 2001US-00770423.

PR 27-JAN-2000; 2000US-0178512P.

XX (GORLACH J.
 ANYV) AN Y.
 HAMI) HAMILTON C M.
 PRIC) PRICE J L.
 RAIN) RAINES T M.
 YUY) YU Y.
 RAME) RAMEKA J G.
 PAGE) PAGE A.
 MATH) MATTHEW A V.
 LEDE) LEDFORD B L.
 WOES) WOESSNER J P.
 HAAS) HAAS W D.
 GARC) GARCIA C A.
 KRIC) KRICKER M.
 SLAT) SLATER T.
 DAVI) DAVIS K R.
 ALLE) ALLEN K.
 HOFF) HOFFMAN N.
 HURB) HURBAN P.

XX Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y, Rameka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD, Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N; Hurban P; DR WPI; 2003-110411/10.

X Novel Arabidopsis thaliana nucleic acid useful for identifying homologous
 PT or related genes, and to create genetically modified and transgenic
 PT organisms, such as plant cells and plants.

XX Claim 1; SEQ ID NO 146; 43pp; English.

The invention describes an Arabidopsis thaliana nucleic acid (I). The
 CC polypeptide (II) encoded by (I), transgenic plant (III) or genetically
 modified cell (IV) are useful for screening a candidate agent for its
 biological effect, by combining the candidate agent with (II), (III) or
 CC (IV), and determining the effect of the candidate agent on (III), (III) or
 CC (IV). (I) is useful for identifying homologous or related genes, for
 CC producing compositions that modulate the expression or function of its
 encoded protein, for mapping functional regions of the protein, in
 CC diagnosis, for studying associated physiological pathways, for genetic
 manipulation of cells, preferably plant cells, in screening assays of
 CC various plant strains to determine the strains that are capable of
 CC withstanding a particular disease or environmental stress, for enhancing
 CC or inhibiting production of biosynthetic product in a plant, for
 CC producing polypeptides as probes for the detection of mRNA in biological
 CC samples, to generate additional copies of (I), to generate ribozymes or
 CC oligonucleotides, as single stranded DNA probes or as triple-strand
 CC forming oligonucleotides, and to create genetically modified and
 CC transgenic organisms, such as plant cells and plants. (II) or (III) is
 CC useful for introducing or improving disease resistance and stress
 tolerance in plants, screening biological active agents, e.g.,
 fungicides, insecticides, etc., and for elucidating biochemical pathways.
 CC (III) is useful as crops for their enhanced disease resistance, enhanced
 CC traits of interest, for screening programs as crops which exhibit
 CC enhanced tolerance to environmental stress, or to produce a factor. This
 sequence represents a nucleic acid that may correspond to naturally
 occurring Arabidopsis thaliana expressed sequences. Note: The sequence
 data for this patent did not form part of the printed specification, but
 was obtained in electronic format directly from the US patent office at
 CC seqdata.uspto.gov/sequence.html?docId=999909771423
 XX Sequence 431 BP; 124 A; 77 C; 103 G; 127 T; 0 U; 0 Other;

Query Match 82.5%; Score 19.8; DB 7; Length 431;
 Best Local Similarity 91.3%; Pred. No. 63; Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCACCAAGAGGACCCAGAT 24
 Db 3 67 AACCCACCAAGAGGACCCAAAT 345

RESULT 3

ABZ12691/C
 ID ABZ12691 standard; DNA; 2868 BP.

XX ABZ12691;
 AC
 DT 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 496.

XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.

XX WO200216655-A2.

XX 28-FEB-2002.

PP 24-AUG-2001; 2001WO-US026685.

XX 24-AUG-2000; 2000US-0227866P.

PR 26-JAN-2001; 2001US-0264647P.

PR 22-JUN-2001; 2001US-030011P.

XX (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Harper JF, Kreps J, Wang X, Zhu T;
 XX
 WPI; 2002-304127/34.
 PT Identifying a stress condition to which a plant cell has been exposed and producing plants with increased tolerance to these abiotic stresses.
 XX
 PS Claim 144; SEQ ID NO 496; 577pp + sequence Listing; English.
 XX
 CC The invention relates to identifying a stress condition to which a plant cell has been exposed comprising: (a) contacting nucleic acid representative of expressed polynucleotides in the plant cell with an array or probes representative of the plant cell genome; and (b) detecting a profile of expressed polynucleotides in the plant cell characteristic of a stress response. The method is useful in the production of transgenic plants, cells and seeds and in producing plants with increased tolerance to abiotic stress. The present sequence is that of an Arabidopsis thaliana stress regulated gene (ABZ12196-BBZ17574) used in methods of the invention. Note: The sequence data for this patent is not represented in the printed specification but is based on sequence information supplied to Derwent by the European Patent Office
 CC
 XX Sequence 2868 BP; 871 A; 581 C; 682 G; 734 T; 0 U; 0 Other;
 SQ
 Query Match 82.5%; Score 19.8; DB 6; Length 2868;
 Best Local Similarity 91.3%; Pred. No. 74;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AGCCACCAAGGAGGCCAGAT 24
 Db 694 AACCCACCARAGGAGGCCAAAT 672
 RESULT 4
 ADA68110/c
 ID ADA68110 standard; DNA; 2868 BP.
 XX
 AC ADA68110;
 XX
 DT 20-NOV-2003 (first entry)
 DE Arabidopsis thaliana gene, SEQ ID 373.
 KW Plant; bacterial infection; fungal infection; viral infection; ds.
 XX
 OS Arabidopsis thaliana.
 PN WO2003000898-A1.
 PR 03-JAN-2003.
 XX
 PP 22-JUN-2001; 2001WO-IB001105.
 XX
 PR 22-JUN-2001; 2001WO-IB001105.
 PA (SYGN) SYNGENTA PARTICIPATIONS AG.
 XX
 PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y, Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
 XX
 DR WPI; 2003-175290/17.
 PT Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.
 XX
 PS Claim 6; SEQ ID NO 373; 899pp; English.
 XX
 CC The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1

comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.

XX Sequence 2868 BP; 871 A; 581 C; 682 G; 734 T; 0 U; 0 Other;
 SQ
 Query Match 82.5%; Score 19.8; DB 7; Length 2868;
 Best Local Similarity 91.3%; Pred. No. 74;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 AGCCACCAAGGAGGCCAGAT 24
 Db 694 AACCCACCARAGGAGGCCAAAT 672
 RESULT 5
 AAA5245/c
 ID AAA5245 standard; DNA; 141 BP.
 XX
 AC AAA5245;
 XX
 DT 07-DEC-2000 (first entry)
 DE Eosinophil activating peptide gene 33.
 KW Bosinophil activation; human; allergy; eosinophilia; cancer; inflammation; signalling peptide; cell adhesion peptide; G-protein coupled receptor; secreted protein; cell surface protein; ds.
 XX
 OS Homo sapiens.
 PN WO200032630-A2.
 XX
 PR 08-JUN-2000.
 PA (SEAR) SEARLE & CO G D.
 XX
 PI Dotson SB, Ma X;
 XX
 DR WPI; 2000-465041/40.
 XX
 PT Novel nucleic acids derived from activated eosinophil cells useful for treating allergic diseases such as asthma comprises a specific nucleotide sequence.
 PT
 XX
 PS Claim 1; Page 114; 125pp; English.
 XX
 CC The present invention relates to a number of nucleotide sequences which encode proteins involved in the activation of eosinophils (AAA52611-A52679). Eosinophils are involved in immune reactions, and these genes and their proteins provide possible targets for new drugs to combat diseases such as asthma, allergic rhinitis, atopic dermatitis, eczema, anaphylaxis, allergic bronchopulmonary aspergillosis, psoriasis, emphysema, leukemias, lymphomas, ovarian cancer, pneumonia, immune disorders, some connective tissue disorders, and inflammatory conditions including septic shock, arthritis, nephritis, inflammatory bowel disease and Crohn's disease
 XX Sequence 141 BP; 17 A; 40 C; 49 G; 35 T; 0 U; 0 Other;
 SQ
 Query Match 79.2%; Score 19; DB 3; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC biological sample
 XX
 SQ Sequence 3205 BP; 571 A; 900 C; 1061 G; 673 T; 0 U; 0 Other;
 Query Match 79.2%; Score 19; DB 7; Length 3205;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAGGCCACCAAGAGGACCC 19
 Db 1081 CAGGCCACCAAGAGGACCC 1063

RESULT 10
 AB268462:C
 ID AB268462 standard; DNA; 3205 BP.
 XX
 AC AB268462:
 XX
 DT 22-APR-2003 (first entry)
 DR Nucleotide sequence of human NOV1 variant SNP1.
 XX
 KW Human; NOV1; hexokinase 3; splice variant; metabolic disorder; diabetes;
 obesity; anorexia; cancer; metastatic melanoma; cancer; enzyme;
 neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
 immune disorder; systemic lupus erythematosus; haematoopoietic disorder;
 cirrhosis; transplant; autoimmune disease; allergy; infection; vaccine;
 KW high blood pressure; haemophilia; wasting disorder; gene therapy;
 KW single nucleotide polymorphism; SNP; SE;
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO2003000918-A2.
 XX
 PD 03-JAN-2003.
 XX
 PF 21-JUN-2002; 2002WO-US019650.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 22-JUN-2001; 2001US-0300290P.
 PR 09-AUG-2001; 2001US-0311285P.
 PR 05-OCT-2001; 2001US-0327345P.
 PR 09-OCT-2001; 2001US-0327892P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Grossé WM, Alsbrook JP, Lepley DM, Burgess CE, Bader JS;
 PI Bansal A, Pena CBA, Shimkets RA;
 DR
 XX
 PT New NOV-1 polypeptides and polyribonucleotides, useful for manufacturing a
 medicament for treating or preventing disorders or syndromes, including
 metabolic disorders, cancer, neurodegenerative and autoimmune disorders.
 PT
 XX
 PT New NOV-1 polypeptides and polyribonucleotides, useful for manufacturing a
 medicament for treating or preventing disorders or syndromes, including
 metabolic disorders, cancer, neurodegenerative and autoimmune disorders.
 PT
 XX
 PT New NOV-1 polypeptides and polyribonucleotides, useful for manufacturing a
 medicament for treating or preventing disorders or syndromes, including
 metabolic disorders, cancer, neurodegenerative and autoimmune disorders.
 PT
 XX
 CC The present sequence represents a variant of NOV1, comprising a single
 nucleotide polymorphism (SNP). NOV1 is a member of the hexokinase 3-like
 family. It is an alternative splice variant. The NOV-1 polypeptide,
 nucleic acid or antibody is useful for manufacturing a medicament for
 treating or preventing disorders or syndromes, including metabolic
 disorders (e.g. diabetes, obesity, or anorexia), cancer (e.g. metastatic
 melanoma), cancer-associated disorders, neurodegenerative disorders (e.g.
 Alzheimer's disease, Parkinson's disease), immune disorders (e.g.
 systemic lupus erythematosus), haematopoietic disorders, cirrhosis,
 transplantation, autoimmune disease, allergies, viral/bacterial/parasitic
 infections, high blood pressure, haemophilia, and wasting disorders
 associated with chronic and various cancers. The polypeptides may be used
 as immunogen to produce antibodies as vaccines, and to screen for
 potential agonist and antagonist compounds. The nucleic acids are also
 useful in gene therapy, in chromosome mapping, tissue typing, and in

CC forensic identification of a biological sample
 XX
 SQ Sequence 3205 BP; 571 A; 901 C; 1061 G; 672 T; 0 U; 0 Other;
 Query Match 79.2%; Score 19; DB 7; Length 3205;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CAGGCCACCAAGAGGACCC 19
 Db 1081 CAGGCCACCAAGAGGACCC 1063

RESULT 11
 ADC10127/C
 ID ADC10127 standard; DNA; 3205 BP.
 XX
 AC ADC10127:
 XX
 DT 18-DEC-2003 (first entry)
 DR Human NOVX polypeptide coding sequence SEQ ID NO: 147.
 XX
 KW ds; gene; cytosolic; antidiabetic; anorectic; cerebroprotective;
 neuroprotective; antiinflammatory; gene therapy; antisense therapy;
 thyromimetic; NOVX; pathology; cancer; diabetes; obesity;
 endocrine disorder; CNS disorder; inflammatory disorder;
 KW chromosome mapping; tissue typing; predictive medicine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003000842-A2.
 XX
 PD 03-JAN-2003.
 XX
 PF 04-JUN-2002; 2002WO-US017443.
 XX
 PR 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0255661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 07-JUN-2001; 2001US-0286575P.
 PR 11-JUN-2001; 2001US-0297414P.
 PR 12-JUN-2001; 2001US-0295573P.
 PR 12-JUN-2001; 2001US-0297567P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298518P.
 PR 18-JUN-2001; 2001US-0299133P.
 PR 19-JUN-2001; 2001US-0299230P.
 PR 21-JUN-2001; 2001US-0299499P.
 PR 22-JUN-2001; 2001US-0300177P.
 PR 26-JUN-2001; 2001US-0301883P.
 PR 28-JUN-2001; 2001US-0301530P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 03-JUL-2001; 2001US-0302951P.
 PR 14-JUL-2001; 2001US-0322297P.
 PR 25-SEP-2001; 2001US-0346699P.
 PR 03-DEC-2001; 2001US-0334747P.
 PR 14-DEC-2001; 2001US-0341562P.
 PR 21-FEB-2002; 2001US-0358656P.
 PR 21-FEB-2002; 2001US-0359122P.
 PR 22-FEB-2002; 2001US-0358978P.
 PR 22-FEB-2002; 2001US-0359034P.
 PR 22-FEB-2002; 2001US-0359035P.
 PR 22-FEB-2002; 2001US-0359121P.
 PR 27-FEB-2002; 2001US-0359964P.
 PR 01-MAR-2002; 2001US-0360858P.
 PR 12-MAR-2002; 2001US-0363430P.
 PR 12-MAR-2002; 2001US-0363676P.
 PR 10-APR-2002; 2001US-071346P.
 PR 04-JUN-2002; 2002US-0037944P.

XX (CURA-) CURAGEN CORP.
 PA (BETA-) BETAGENE INC.
 XX
 XX
 PI Agee ML, Anderson DW, Berghs C, Casman SJ, Catterton E;
 PI Dippipo VA, Edinger SR, Eisen A, Ellerman K, Gangolli EA;
 PI Gerlach VL, Gorman L, Guo X, Herrmann JL, Hiltz T, Ji W, Kekuda R;
 PI Khramtsov NV, Li L, Liu X, Malyankar UM, Miller CE, Millet I;
 PI Ort T, Padigaru M, Patturajai M, Pena CEA, Rastelli L, Rieger DK;
 PI Rothenberg ME, Shenvoy SG, Shimkets RA, Smithson G, Spaderna SK;
 PI Spytek KA, Stone DJ, Vernet CAM, Zhong H, Zhong M, Alsobrook JP;
 PI Burgess CB, Lepley DM;
 XX DR WPI; 2003-210149/20.
 DR P-PSDB; ADC10128.
 XX
 PT New isolated NOX polypeptides and nucleic acid molecules useful for
 PT preventing, preventing and diagnosing pathological conditions with NOX-
 PT associated disorders, such as cancer, obesity, diabetes and inflammatory
 PT diseases.
 XX
 PS Claim 20; SEQ ID NO 147; 772pp; English.
 XX The invention relates to novel isolated polypeptides, mature form of the
 CC polypeptide, a sequence that is 95% identical to the polypeptide or the
 CC polypeptide comprising one or more conservative substitutions. The NOX
 CC polypeptide is useful for treating or preventing a pathology associated
 CC with the polypeptide e.g. disorders associated with aberrant expression
 CC or activity of the polypeptide, such as cancer, diabetes, obesity, and
 CC endocrine, CNS and inflammatory disorders. They can also be used in
 CC various detection and screening assays, chromosome mapping, tissue typing
 CC and predictive medicine. This sequence corresponds to the coding sequence
 CC for one of the polypeptides of the invention.
 XX Sequence 3205 BP; 571 A; 900 C; 1061 G; 673 T; 0 U; 0 Other;
 CC Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 CC Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy Oy
 1 CAGGCCACCAAAAGGAGGCC 19
 Db 1081 CAGCCACCAAAAGGAGGCC 1063
 RESULT 12
 AAV00090/C
 ID AAV00090 standard; cDNA; 3692 BP.
 AC AAV00090;
 XX
 DT 17-MAR-1998 (first entry)
 DB Rat hexokinase III encoding cDNA.
 XX
 KW Yeast; trehalose-6-phosphate synthase; hexokinase; ribozyme; stimulator;
 KW insulin; glucose; diabetes; ss.
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT 81..2855 /*tag= a
 FT /product= "Hexokinase_III"
 PN WO9726357-A1.
 PD 24-JUL-1997.
 XX
 PR 17-JAN-1997; 97WO-US000787.
 XX
 PR 19-JAN-1996; 96US-00588983.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (BETA-) BETAGENE INC.
 PA (UNMS) UNTV MICHIGAN STATE.
 XX
 PA Newgard CB, Han H, Becker TC, Wilson JE;
 PA (TEXA) UNIV TEXAS SYSTEM.

PA (BETA-) BETAGENE INC.
 XX
 PI Newgard CB, Han H, Thiagpen AE, Normington KD;
 XX
 PI WPI; 1997-385343/35.
 DR P-PSDB; AW37430.
 XX
 PT Mammalian cell with reduced activity of low Km hexokinase - caused by
 PT specific ribozyme or stimulator of trehalose-6-phosphate production,
 PT particularly for in vitro or in vivo insulin production.
 XX
 PS Disclosure; Page 195-198; 263pp; English.
 CC A mammalian cell has been developed comprising an effective amount of a
 CC low Km hexokinase-inhibitor selected from: (a) an agent that stimulates
 CC production of trehalose-6-phosphate (TCP); or (b) a ribozyme specific for
 CC low Km hexokinase. The present sequence encodes rat hexokinase III. The
 CC cells are particularly used to produce insulin, in response to glucose or
 CC other secretagogues, either in vitro or in vivo (for treating diabetes),
 CC but may also be used to produce many other therapeutic proteins, e.g.
 CC from a gene under control of the insulin promoter which is therefore
 CC responsive to glucose. Growth-inhibited cells (having altered hexokinase
 CC activity) are used to produce proteins, e.g., insulin or antibodies, in
 CC vivo or in vitro. Reduction in low Km hexokinase activity provides cells
 CC in which insulin secretion is induced at glucose concentrations closer to
 CC the normal range than in the parent cell (nearly homeostatic secretion).
 CC Implanted cells of reduced low Km hexokinase activity are expected to
 CC survive longer in the host
 XX Sequence 3692 BP; 821 A; 969 C; 1112 G; 790 T; 0 U; 0 Other;
 CC Best Local Similarity 79.2%; Pred. No. 1.7e+02;
 CC Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy Oy
 1 CAGGCCACCAAAAGGAGGCC 19
 Db 1087 CAGCCACCAAAAGGAGGCC 1069

RESULT 13
 AAV00160/C
 ID AAV00160 standard; cDNA; 3692 BP.
 XX
 AC AAV00160;
 XX
 DT 17-MAR-1998 (first entry)
 XX
 DE Rat hexokinase III encoding cDNA.
 XX
 KW Hexokinase; competitive enzyme; inhibitor; insulin; diabetes; B cell;
 KW glucose; glucokinase; ss.
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT 81..2855 /*tag= a
 FT /product= "Hexokinase_III"
 PN WO9726322-A2.
 XX
 PR 24-JUL-1997.
 XX
 PR 17-JAN-1997; 97WO-US000786.
 XX
 PR 19-JAN-1996; 96US-00588976.
 XX
 PA (TEXA) UNIV TEXAS SYSTEM.
 PA (BETA-) BETAGENE INC.
 PA (UNMS) UNTV MICHIGAN STATE.
 XX
 PA Newgard CB, Han H, Becker TC, Wilson JE;

XX WPI; 1997-385327/35.
 DR PR
 DR XX (P-PSDB; AAW37443).
 XX PT Mammalian cell with reduced activity of low Km hexokinase - caused by PT in vitro or in vivo insulin production.
 XX PS Disclosure; Page 195-198; 254pp; English.
 XX CC A mammalian cell has been developed which includes enough of a competitive inhibitor of low Km (Michaelis constant) hexokinase to reduce the hexokinase activity in the cell. The present sequence encodes rat hexokinase III. The cells are particularly used to produce insulin, in response to glucose or other secretagogues, either in vitro or in vivo (for treating diabetes), but may also be used to produce many other therapeutic proteins, e.g. from a gene under control of the insulin promoter and thus responsive to glucose. Inhibition of the inhibitor can be used more generally to reduce the growth rate of cells, e.g. CC CC CC Sequence 3692 BP; 821 A; 969 C; 1112 G; 790 T; 0 U; 0 Other;
 XX SQ Best Local Similarity 100.0%; Pred. No. 1.7e+02; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CAGGCCCAAGAGGACCC 19
 Db 1087 CAGGCCCAAGAGGACCC 1069

RESULT 14 ABT42065/C
 ID ABT42065 standard; DNA; 3692 BP.
 XX AC ABT42065;
 XX DT 26-JUN-2003 (first entry)
 XX DB Toxicity modelling related rat gene SEQ ID No 1767.
 KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
 KW database; drug screening; toxicity assay; rat; da.
 OS Rattus norvegicus.
 PN WO200295000-A2.
 XX PD 28-NOV-2002.
 XX PP 22-MAY-2002; 2002WO-US016173.
 XX PR 22-MAY-2001; 2001US-0292335P.
 PR 13-JUN-2001; 2001US-029523P.
 PR 13-JUN-2001; 2001US-0298925P.
 PR 10-JUL-2001; 2001US-0303807P.
 PR 10-JUL-2001; 2001US-0303910P.
 PR 28-APR-2001; 2001US-0315047P.
 PR 27-SEP-2001; 2001US-0324928P.
 PR 22-OCT-2001; 2001US-0304622P.
 PR 01-NOV-2001; 2001US-0305067P.
 PR 21-NOV-2001; 2001US-0333805P.
 PR 06-DEC-2001; 2001US-036144P.
 PR 19-DEC-2001; 2001US-0340873P.
 PR 21-FEB-2002; 2002US-0357842P.
 PR 21-FEB-2002; 2002US-0357844P.
 PR 21-FEB-2002; 2002US-0364134P.
 PR 15-MAR-2002; 2002US-0364134P.
 PR 08-APR-2002; 2002US-0370144P.
 PR 08-APR-2002; 2002US-0370206P.
 PR 08-APR-2002; 2002US-0370247P.

RESULT 14 ABT42065/C
 ID ABT42065 standard; DNA; 3692 BP.
 XX AC ABT42065;
 XX DT 26-JUN-2003 (first entry)
 XX DB Toxicity modelling related rat gene SEQ ID No 1767.
 KW Toxic effect; gene expression profile; renal toxicity; toxicity marker;
 KW database; drug screening; toxicity assay; rat; da.
 OS Rattus norvegicus.
 PN WO200295000-A2.
 XX PD 28-NOV-2002.
 XX PP 22-MAY-2002; 2002WO-US016173.
 XX PR 22-MAY-2001; 2001US-0292335P.
 PR 13-JUN-2001; 2001US-029523P.
 PR 13-JUN-2001; 2001US-0298925P.
 PR 10-JUL-2001; 2001US-0303807P.
 PR 10-JUL-2001; 2001US-0303910P.
 PR 28-APR-2001; 2001US-0315047P.
 PR 27-SEP-2001; 2001US-0324928P.
 PR 22-OCT-2001; 2001US-0304622P.
 PR 01-NOV-2001; 2001US-0305067P.
 PR 21-NOV-2001; 2001US-0333805P.
 PR 06-DEC-2001; 2001US-036144P.
 PR 19-DEC-2001; 2001US-0340873P.
 PR 21-FEB-2002; 2002US-0357842P.
 PR 21-FEB-2002; 2002US-0357844P.
 PR 21-FEB-2002; 2002US-0364134P.
 PR 15-MAR-2002; 2002US-0364134P.
 PR 08-APR-2002; 2002US-0370144P.
 PR 08-APR-2002; 2002US-0370206P.
 PR 08-APR-2002; 2002US-0370247P.

RESULT 15 ADBB8419/C
 ID ADBB8419 standard; DNA; 3692 BP.
 XX AC ADBB8419;
 XX DT 04-DEC-2003 (first entry)
 XX DB Toxicity-related gene, SEQ ID 3445.
 KW Toxic; toxin; gene expression profile; hepatotoxicity; liver;
 KW drug screening; toxicity assay; da.
 OS Unidentified.
 PN WO200306624-A2.
 XX PD 07-AUG-2003.
 XX PP 31-JAN-2003; 2003WO-US003194.
 XX PR 31-JAN-2002; 2002US-00060087.
 PR 15-MAR-2002; 2002US-0361045P.
 PR 15-MAR-2002; 2002US-0361045P.
 PR 30-DEC-2002; 2002US-0436643P.
 XX PA (GENE-) GENE LOGIC INC.
 XX PI Mendlrick D, Porter M, Johnson K, Higgs B, Castle A, Blashoff M;

PR 17-APR-2002; 2002US-0372794P.
 PR 21-APR-2002; 2002US-0371679P.

XX PA (GENE-) GENE LOGIC INC.

XX PI Mendlrick D, Porter M, Johnson K, Higgs B, Castle A, Blashoff M;

XX PR WPI; 2003-148464/14.

XX PT Predicting at least one toxic effect of a compound, useful for toxicity modeling, comprises preparing a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database.

XX PS Example 4; Page; 446pp; English.

XX CC The invention relates to a novel method of predicting at least one toxic effect of a compound. The method comprises a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database comprising at least part of the data or information given in the specification. The methods are useful for predicting at least one toxic effect of a compound, predicting the progression of a toxic effect of a compound, predicting the renal toxicity of a compound, or identifying toxicity markers in tissues or cells exposed to known renal toxin. The genes are useful as toxicity markers in drug screening and toxicity assays, in monitoring disease or physiological states, or disease progression. This polynucleotide represents a rat DNA sequence relating to the toxic effect database described in the specification. NOTE: The sequence data for this patient did not form part of the printed specification, but was obtained in electronic format directly from the World Intellectual Property Organization.

XX SQ Sequence 3692 BP; 821 A; 969 C; 1112 G; 790 T; 0 U; 0 Other; Best Local Similarity 100.0%; Score 19; DB 7; Length 3692; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGGCCCAAGAGGACCC 19
 Db 1087 CAGGCCCAAGAGGACCC 1069

XX
DR
XX

WPI; 2003-689530/65.

PT Predicting a toxic effect of a compound, useful in identifying toxicity assays,
PT markers in liver tissues or cells for drug screening and toxicity assays,
PT comprises preparing gene expression profile of tissue or cells exposed to
PT the compound.

XX
PS Claim 1; SEQ ID NO 3445; 1156pp; English.

CC The present invention relates to a method for predicting a toxic effect
CC of a compound. The method comprises preparing a gene expression profile
CC of a tissue or cell sample exposed to the compound and comparing the
CC gene expression profile to a database comprising SEQ ID 1-4925, where
CC differential expression of the gene indicates at least one toxic effect.
CC The method is useful for predicting at least one toxic effect of a
CC compound, predicting hepatotoxicity or the progression of a toxic effect
CC of a compound, identifying an agent that modulates the onset or
CC progression of a toxic response, predicting the cellular pathways that a
CC compound modulates in a cell, and identifying an agent that modulates at
least one activity of a protein. The method and compositions of the
CC present invention using a database of genes having liver toxin-induced
CC differential expression, are useful in identifying toxicity markers in
CC liver tissues or cells for drug screening and toxicity assays. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at [ftp://www.wipo.int/pdb/published_pct_sequences](http://www.wipo.int/pdb/published_pct_sequences).

XX
SQ Sequence 3692 BP; 821 A; 969 C; 1112 G; 790 T; 0 U; 0 Other;

Query Match 79.2%; Score 19; DB 9; Length 3692;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY ||||| CAAGCCACCAAGAGGACCC 19
Db 1087 CAGCCACCAAGAGGACCC 1069

Search completed: July 30, 2004, 20:33:04
Job time : 386 secs



GenCore version 5.1.6
(c) 1993 - 2004 Compugen Ltd.

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Run on: July 30, 2004, 19:11:31 ; Search time 2503 Seconds
(without alignments)
415.594 Million cell updates/sec

Title:

US-09-831-272-7

Perfect score:

24

Sequence:

1 cagccacccaaggagccccagaat 24

Scoring table:

IDENTY_NUC

Gapop 10⁻⁰, Gapext 1.0

Maximum DB seq length: 0

Minimum DB seq length: 0

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

OM nucleic - nucleic search, using sw model

Copyright (c) 1993 - 2004 Compugen Ltd.

RESULT 1

Result No.	Score	Query Match	Length	DB ID	Description
1	24	100.0	24	6 AX025607	AX025607 Sequence
2	24	100.0	3549	8 AF239835	AF239835 Petroselinum
3	19.8	82.5	2868	6 AX505801	AX505801 Sequence
4	19.8	82.5	2868	6 AX51563	AX51563 Sequence
5	19.8	82.5	2962	8 AY133782	AY133782 Arabidopsis
6	19.8	82.5	3107	8 AF081202	AF081202 Arabidopsis
7	19.8	82.5	3295	8 AF088901	AF088901 Lithium lo
8	19.8	82.5	3383	8 AY080601	AY080601 Arabidopsis
9	19.8	82.5	81843	10 AL672309	AL672309 Mouse DNA
10	19.8	82.5	91071	8 AF02339	AF02339 Arabidopsis
11	19.4	80.8	813	11 BV011807	BV011807 S21286764
12	19.4	80.8	173763	2 AC134552	AC134552 Mus muscu
13	19.4	80.8	18334	10 AC017416	AC017416 Mus muscu
14	19.4	80.8	213625	10 AL805918	AL805918 Mouse DNA
15	19.2	80.0	1603	8 VCCYNA	X69193 V-cardunculus
16	19.2	80.0	1970	8 AF082029	AF082029 Hemimocula
17	19.2	80.0	130709	2 AC102945	AC102945 Mus muscu
18	19.2	80.0	189891	2 AC119058	AC119058 Papio anu
19	19.2	80.0	19132	2 AC16895	AC16895 Rattus no
20	19.2	80.0	19082	10 AL591490	AL591490 Mouse DNA
21	19.2	80.0	206145	2 BX813306	BX813306 Mus muscu
22	19.2	80.0	214393	2 BX517302	BX517302 Mouse DNA
23	19.2	80.0	247299	2 AC112220	AC112220 Rattus no
24	19.2	80.0	247202	2 AC073807	AC073807 Mus muscu
25	19.2	80.0	255871	2 AC095418	AC095418 Rattus no
26	19.2	80.0	263367	2 AC099427	AC099427 Rattus no
27	19.2	80.0	260145	2 AX410736	AX410736 Sequence
28	19.2	80.0	214393	2 AC112220	AC112220 Rattus no
29	19.2	80.0	247299	2 AC073807	AC073807 Mus muscu
30	19.2	80.0	247202	2 AC095418	AC095418 Rattus no
31	19.2	80.0	255871	2 AC099427	AC099427 Rattus no
32	19.2	80.0	263367	2 AC099427	AC099427 Rattus no
33	19.2	80.0	260145	2 AX410736	AX410736 Sequence
34	19.2	80.0	214393	2 AC095418	AC095418 Rattus no
35	19.2	80.0	247299	2 AC073807	AC073807 Sequence
36	19.2	80.0	247202	2 AC112220	AC112220 Rattus no
37	19.2	80.0	255871	2 AC095418	AC095418 Rattus no
38	19.2	80.0	263367	2 AC099427	AC099427 Rattus no
39	19.2	80.0	260145	2 AX410736	AX410736 Sequence
40	18.8	78.3	400	11 G17829	G17829 human STR S
41	18.8	78.3	2228	3 AV069309	AV069309 Drosophila
42	18.8	78.3	2452	5 BC044215	BC044215 Xenopus l
43	18.8	78.3	2999	8 AY128932	AY128932 Arabidopsis
44	18.8	78.3	3335	8 AF081203	AF081203 Arabidopsis
45	18.8	78.3	3548	8 AY093052	AY093052 Arabidopsis
46	18.8	78.3	73181	9 AC117455	AC117455 Homo sapi

ALIGNMENTS

TITLE Chimeric promoters capable of mediating gene expression in plants upon pathogen infection and uses thereof
 JOURNAL Patent: WO 029592-A 7-25-MAY-2000;
 MAX PLANCK GESELLSCHAFT (DE); LOEGMANN ELIKE (DE); SOMSICH IMRE (DE); HAHLBROCK KLAUS (DE); KIRSCH CHRISTOPH (DE); RUSHTON PAUL (GB)

FEATURES source
 (GB) Location/Qualifiers
 1. -24
 /organism="Petroselinum crispum"
 /mol_type="unassigned DNA"
 /db_xref="taxon:4043"

ORIGIN Query Match 100.0%; Score 24; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0;
 Matches 24; Conservative 0; Indels 0; Gaps 0;

RESULT 2 Qy 1 CAGGCCACCAAGAGGACCCGAAAT 24
 Db 1 CAGGCCACCAAGAGGACCCGAAAT 24

ACCESSION AP229835
 VERSION AP239835.1 GI:10945374
 KEYWORDS petroselinum crispum (parsley)
 ORGANISM Petroselinum crispum
 REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; asterids; campanulids; Apiales; Apioideae; Apioideae; apiod
 AUTHORS Kirsch,C., Takamiya-Wik,M., Schmelzer,E., Hahlbrock,K. and Somsich,I.E.

TITLE A novel regulatory element involved in rapid activation of parsley ELT7 gene family members by fungal elicitor or pathogen infection
 JOURNAL Mol. Plant Pathol. 1 (4), 243-251 (2000)
 REFERENCE 1 (bases 1 to 3549)
 2 (bases 1 to 3549)
 AUTHORS Kirsch,C., Takamiya-Wik,M. and Somsich,I.E.
 TITLE Direct Submission
 JOURNAL Submitted (29-FEB-2000) Biochemistry, Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne Weg 10, Koeln, NRW D-50829, Germany

FEATURES source
 1. 3549 Location/Qualifiers
 /organism="Petroselinum crispum"
 /db_xref="taxon:4043"
 /note="from cultured cells"
 gene 380 .1782
 /gene="ELT7.1"
 380 .1782
 /gene="ELT7.1"
 /product="fatty acid desaturase/hydroxylase-like protein
 ELT7.1"
 551 .1702
 /gene="ELT7.1"
 /codon_start:1
 /product="fatty acid desaturase/hydroxylase-like protein
 ELT7.1"
 /protein_id="PAG2521.1"
 /db_xref="GI:10945375"
 /translation="MGAGCGRMSAASNKQSAEARRAPRHEKPPPTIGDLKAIPAH
 FEKLITPPRLYDQMLAYALVATNTINQIPLPHPNVLYGAAVQCGVLTGAM
 VVGEBCDDIAFDSDGIVNDLVALVHSSMPPFSMKSHRRHANTOSLENDTGG
 RFRSNRNYKILNNPPGRVLVWITLUVGPPLVLMFNVSGKVERMTSHYDPHSPLY

ORIGIN Query Match 100.0%; Score 24; DB 6; Length 24;
 Best Local Similarity 100.0%; Pred. No. 0.31; Mismatches 0;
 Matches 24; Conservative 0; Indels 0; Gaps 0;

RESULT 2 Qy 1 CAGGCCACCAAGAGGACCCGAAAT 24
 Db 1 CAGGCCACCAAGAGGACCCGAAAT 24

ACCESSION AP229835
 VERSION AP239835.1 GI:10945374
 KEYWORDS petroselinum crispum (parsley)
 ORGANISM Petroselinum crispum
 REFERENCE Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; asterids; eurosids II; Brassicales; Brassicaceae; Arabidopsis
 AUTHORS Harper,J.F., Kreps,J., Wang,X. and Zhu,T.
 TITLE Stress-regulated genes of plants, transgenic plants containing same, and methods of use
 JOURNAL Patent: WO 0216655-A 495 28-FEB-2002;
 The Scripts Research Institute (US); Syngenta Participations AG (CH)

FEATURES source
 1. 2868 Location/Qualifiers
 /organism="arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN Query Match 82.5%; Score 19.8; DB 6; Length 2868;
 Best Local Similarity 91.3%; Pred. No. 52; Mismatches 21; Conservative 0; Indels 2; Gaps 0;

RESULT 4 Qy 2 AGCCACCAAGAGGACCCGAAAT 24
 Db 694 AACCCACCAAGAGGACCCGAAAT 672

ACCESSION AX651563
 DEFINITION Sequence 373 from Patent WO03000898.
 VERSION AX651563
 KEYWORDS SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicots; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,

Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.

TITLE plant genes involved in defense against pathogens

PATENT: WO 03000898-A 373 03-JAN-2003;

Syngenta Participations AG (CH)

FEATURES Location/Qualifiers

ORIGIN SOURCE 1. . 2868

/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

RESULT 5 Query Match

Best Local Similarity 82.5%; Score 19.8; DB 6; Length 2868;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCACCAAGGAGGACAGAT 24

Db 694 AACCAACGAGGACCAATT 672

DEFINITION Arabidopsis thaliana clone U11765 putative villin 2 protein
 ACCESSION AT1G3782.1 GI:22136973
 VERSION FJ11CDNA
 KEYWORDS SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCES 1 (bases 1 to 2962)

AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Wong, C., Wu, H.C., Yu, G., Yuan, S., Bowser, L., Carninci, P., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

TITLE Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2962)

AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Deng, J.M., Hsuan, V.W., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Wallender, E.K., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

COMMENT Direct Submission

JOURNAL Submitted (17-JUL-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of R AFLP cDNAs (R AFLP cDNA : 'RIKEN Arabidopsis Full Length cDNA'): Seki, M., Naruseka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

ORIGIN

Query Match

Best Local Similarity 91.3%; Score 19.8; DB 8; Length 2962;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCACCAAGGAGGACAGAT 24

Db 694 AACCAACGAGGACCAATT 672

RESULT 6 Query Match

Best Local Similarity 82.5%; Score 19.8; DB 8; Length 2962;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCACCAAGGAGGACAGAT 24

Db 694 AACCAACGAGGACCAATT 672

DEFINITION Arabidopsis thaliana (thale cress)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicots; core eudicots;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 3107)

AUTHORS Klaire, U., Friedrich, E., Kost, B., Louvard, D. and Chua, N.H.

TITLE Villin-like actin-binding proteins are expressed ubiquitously in Arabidopsis

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

		Plant Physiol. 122 (1), 35-48 (2000)	Vidalii, L., Yokota, E., Cheung, A., Wu, H.M., Shimmen, T. and Hepler, P.K.
		JOURNAL MEDLINE PUBMED REFERENCE	Submitted (31-AUG-1998) Biology, University of Massachusetts, Morrill Science Center, Amherst, MA 01003, USA
		JOURNAL	Klahre, U. and Chua, N.-H.
		TITLE	Direct Submission
		ACCESSION	10631247
		VERSION	2 (bases 1 to 3107)
		KEYWORDS	
		ORGANISM	Lilium longiflorum (trumpet lily)
		SOURCE	Lilium longiflorum (trumpet lily)
		REFERENCE	1 (bases 1 to 3295)
		AUTHORS	Viddali, L., Yokota, E., Cheung, A., Shimmen, T. and Hepler, P.K.
		TITLE	The 135 kDa actin-binding protein from <i>Lilium longiflorum</i> pollen is the plant homolog of villin
		JOURNAL	Protoplasma 209, 283-291 (1999)
		REFERENCE	2 (bases 1 to 3295)
FEATURES	SOURCE	JOURNAL	AUTHORS
		Rockefeller University, New York Avenue, New York, NY 10021, USA	Vidalii, L., Yokota, E., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashi, Kawai, Y., Iwai, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Mianda, M., Naruseka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
		Location/Qualifiers	
1.	3107	/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/cultivar="Landsberg"	
		/db_xref="taxon:3702"	
		/chromosome="T1"	
		gene	1. - 3107
		CDS	1. - 2931
		gene	/gene="VIN2"
		gene	/note="Atvillin; similar to protein encoded by T11A7.16
		gene	/codon_start_1
		gene	/product="villin 2"
		gene	/protein_id="PAC11606_1"
		gene	/db_xref="GI:3415115"
		gene	/translation="MSTKVLDPARQAGQKPGTEIWRITENFEAVPVIREPKHSEHCKYMGD
		gene	TYVLTQTTGKGGATYRDLPHWKGQSDQBACTAAVTCVLEDAVLGGRAYVORHVIQG
		gene	HEEDKFLSYFFCILCLERGGVASSGFKTVBEEVTRFLYCKGRKRAIRKQVPPRSSL
		gene	NHDQFVFLDTERKYYQFNQANSNQOERAKALEVQYVQFLDKYHGTGCVAVDVGKLDT
		gene	ESISGAFAWFLFGGAPRKVKAANDDVPESTPPKLGKTSKASSETTRDPLF
		gene	TKCYLUDCGATIVWGRGTVQDERKAASOSAEFLASBNRPGKATHYTRVQGSHS
		gene	EKSNFDSPHSGSAVPGNGEGRGVKAALIKQOGVKGKJAKSAVNPEDIPPLIESGGKL
		gene	TYFLANTMSNSLKGPGVQGRVKEGRKPOFVALQPMVLUKGQGSSKGMSSES
		gene	TDETYTPSPASIALVQVSGVHNKAQVETVATSAINSNCFLLQSGTSMFLHNGST
		gene	HBDQELATKVAETKQVSGTILKHKRGESTSTEFLALSGKTSKASSETTRDPLF
		gene	SPAFRNGKQEVETYNPAQDLITDYLTHAEVFWGQCVPKEKOTWFGK
		gene	YDLAGSLIGLHRPVPKINEGNECPCTTYSWDAKATVQGNSNQPKQASLIFGH
		gene	HYVDEDKSINGNQOCRUPVQASODRNEHSIDGPQRORABAAL
		gene	ALSSAFNSSTSSTSPPPNGTQSOSARAAMVAALSOVLYENKNSPDSPTRST
		gene	SSNAADDIPLTEAMDEESEVAEGLAEKEREKSVSPADETEAKQETEQGDSEIOPSG
		gene	ATPMEQJRAKSNSNPVTCIDFERRAVYISERECVSRCERKAFPNNPFRWKDILKCK
		gene	FDLFI"
ORIGIN			
RESULT 7		Query Match	82.5%; Score 19.8; DB 8; Length 3107;
		Best Local Similarity	91.3%; Pred. No. 52;
		Matches	21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	2	AGCCACCAAAAGGAGCCAGAT	24
Db	694	AACACCAAAAGGAGCCAAAT	672
ORIGIN			
RESULT 8		Query Match	82.5%; Score 19.8; DB 8; Length 3295;
		Best Local Similarity	91.3%; Pred. No. 52;
		Matches	21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	2	AGCCACCAAAAGGAGCCAGAT	24
Db	822	AGCCACCAAAAGGAGCCAAAT	800
ORIGIN			
RESULT 8		Query Match	82.5%; Score 19.8; DB 8; Length 3295;
		LOCUS	AY080601
		DEFINITION	Arabidopsis thaliana putative villin 2 protein (At2g41740) mRNA,
		VERSION	complete cds.
		VERSION	AY080601
		KEYWORDS	
		ORGANISM	Arabidopsis thaliana (thale cress)
		SOURCE	Arabidopsis thaliana
		REFERENCE	1 (bases 1 to 3383)
		AUTHORS	Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashi, Kawai, Y., Iwai, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Mianda, M., Naruseka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K.,
		TITLE	
		JOURNAL	
		REFERENCE	

Davis,R.W., Ecker,J.R. and Theologis,A.	TITLE	ATFMYEQLRAKSENPTVGTIDKRREAYLSEERQSVFGIEKEAFNNLPRWKQDLKKK
Arabidopsis Full Length cDNA Clones	JOURNAL	FDI,F"
Unpublished	REFERENCE	/gene="At2g41740"
2 (bases 1 to 3383)	AUTHORS	
Yanada,K., Banh,J., Chan,M.M., Chang,C.H., Chang,E., Dale,J.M.,		
Deng,J.M., Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L.,		
Tang,C.C., Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L.,		
Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J.,		
Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.J.,		
Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,		
Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A.,		
Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.		
Direct Submission	TITLE	
Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan	JOURNAL	
Street, Albany, CA 94710, USA	COMMENT	
RIKEN Genomic Sciences Center (GSC) members carried out the		
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN		
Arabidopsis Full Length cDNA'); Seki,M., Narusaka,M., Ishida,J.,		
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,		
Hayashizaki,Y. and Shinozaki,K.		
The Salk, Stanford, PGSC (SSP) Consortium members carried out the		
sequencing and annotation of the RAPL cDNAs: Yanada,K., Banh,J.,		
Chen,M.M., Chang,C.H., Chang,E., Dale,J.M., Deng,J.M.,		
Goldsmith,A.D., Lee,J.M., Onodera,C.S., Quach,H.L., Tang,C.C.,		
Toriumi,M., Wu,H.C., Yamamura,Y., Yu,G., Bowser,L., Chen,H.,		
Chauk,R., Jones,T., Karlin-Neumann,G., Kim,C., Lam,B., Lin,J.,		
Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P.,		
Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
Yanada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to		
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC)		
contributed equally to this work as PIs.		
Annotation is based on the January 2002 version of the Arabidopsis		
genome submitted to Genbank.		
Location/Qualifiers	FEATURES	
source		
1..3383		
/organism="Arabidopsis thaliana"		
/mol_type="mRNA"		
/db_xref="Taxon:3702"		
/chromosome="2"		
/clone="RPL06-70-H22 (R11765)"		
/ecotype="Columbia"		
/note="This clone is in a modified pBluescript vector		
(Lambda ZAP) as a xho1/Sst1 insert."		
1..3383		
/gene="At2g41740"		
1..203		
/gene="At2g41740"		
204..3134		
/gene="At2g41740"		
/codon_start=1		
/evidence=experimental		
/product="putative villin 2 protein"		
/protein_id="AAI8502.1"		
/db_xref="GI:19310538"		
/translation="MSTKVLPDPFOCAGOKRGTEWRIENFEAVPVPKSEHGFYNGD		
TYVLTQTONKGAVYDIIHWIGKTSQEBAGTAVKTFELDAVIGGRAVOHREOG		
HSDKPSYKPCPDTLEGGAGKGVVEETVYTKGRRAIRKQVPLRASSL		
NHDVFLTDLTTRKVKYOFQGANMSNTORAKALEVQVTLKDQHGTCDAVYDDGKLST		
ESDGAHWLFGGABTRGKVNADDIVPESTPPKLYCTDGKMPEDIGLSMLEN		
TKCYLUDGCGAEEYIWGRVTOQDVERKASQSKAEFLASENPKATEVTRIQGYSHS		
FKSNFSPWSSESATPENBEGKVAALNPKVQKQGNGKQJAKSAVPNEDIPPLESGKL		
EWVWYKVKYKPLKEDIGKLYGDCYLVYTHSCKRDYEVLSWFGKKSIPEDQ		
TAIRLANTMSNLKGRGVQGRYLGKEPPQVVALPNUVKGLOSSYKESBES		
HEQELATKYLELKPCITKHKEGSTESPFWAFLGGKONTSMFLWINGNST		
SFAFNRRKFQFEEVINFQADDLIDTQHDFWVQGCVVERKQTVNEFCQ		
YVLLACSLBEGLKHVKVPGKNCERCFDTYFSWDAKTVQWGNFSQKKSLRGEH		
HVEDKNSGNGCQKRBAAELAALANSAFNNSRNPRASSDRLNHSHDGRQRARLA		
ALSAFNSSSSSKSPPPRPGTQSQRARAHVALSQUVVAENKSPPSPTSRARLA		
SSNPADDIPLTEAKDDEBASEEVAGLEAKEEEBVSPPADETEAKOTEEQGDSEIOPSG		
3' UTR	FEATURES	
	source	
1..81843		
/organism="Mus musculus"		
/mol_type="genomic DNA"		
/db_xref="Taxon:10090"		
/chromosome="X"		
/clone="RP23-466J17"		
/clone_lid="RPCI-23"		
ORIGIN	FEATURES	
Query Match	source	82.5%; Score 19.8; DB 10; Length 81843;
Best Local Similarity		91.3%; Pred. No. 57;
Matches		0; Mismatches 2; Indels 0; Gaps 0;
QY	1	CAGCCACCAAGAGGACCCAGAA 23
Db	21198	CAGCCACCAAGAGGACCCAGAA 21220

JOURNAL Submitted (27-SEP-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 173763)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 COMMENT Submitted (20-MAR-2003) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 On Mar 20, 2003 this sequence version replaced gi:29124268.

Center: Washington University Genome Sequencing Center
 Web site: <http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@watsont.wustl.edu
 Project Information
 Center project name: M_BB0279K10

----- Summary Statistics -----

Sequencing vector: M13; 0%
 Sequencing vector: plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.99319
 Consensus quality: 176608 bases at least Q40
 Consensus quality: 173012 bases at least Q30
 Consensus quality: 173325 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 2809: contig of 209 bp in length
 * 2810 2909: gap of unknown length
 * 2910 2614: contig of 2237 bp in length
 * 2614 2626: gap of unknown length
 * 2624 2626: contig of 63584 bp in length
 * 89830 89930: gap of unknown length
 * 89931 89933: contig of 83833 bp in length.

FEATURES

source

1. e173763
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="UNK"
 /clone="RP24-279K10"
 1..2809
 misc_feature
 /note="assembly_name:Contig23"
 2910 ..26146
 /note="assembly_name:Contig24"
 misc_feature
 /note="assembly_name:Contig25"
 26247 ..89830
 /note="assembly_name:Contig26"
 89931 ..173763
 /note="assembly_name:Contig26"

ORIGIN

Query Match 80.8%; Score 19.4; DB 2; Length 173763;
 Best Local Similarity 95.2%; Pred. No. 93; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CAGCCACCAAGGAGCCAG 21
 Db 91787 CGGCCACCAAGGAGCCAG 91807

RESULT 13

AC087416 AC087416 183349 bp DNA linear ROD 15-NOV-2002
 LOCUS Mus musculus chromosome 2 clone rp23-224116 strain C57BL/6J,
 DEFINITION Mus musculus chromosome 2 clone rp23-224116 strain C57BL/6J,
 complete sequence.

REFERENCE AC087416 AC087416.18 GI:2501342
 VERSION
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Mammalia; Butcheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Mus musculus BAC Clone rp23-224116
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (01-JAN-2001) Department of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 3 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUL-2001) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 4 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUL-2001) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 5 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (14-JUL-2001) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 6 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (01-AUG-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 REFERENCE 7 (bases 1 to 183349)
 AUTHORS Kim,J., Shauill,S., Yao,Z. and Roe,B.A.
 TITLE Direct Submission
 JOURNAL Submitted (15-NOV-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
 COMMENT On Nov 15, 2002 this sequence version replaced gi:22002184.
 ----- Genome Center
 Center: Department of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code:UOKONG

FEATURES

source

1..183349
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /chromosome="2"
 /clone="RP23-224116"
 /clone_lib="RPCI - 23 Female (C57BL/6J) Mouse BAC Library"

ORIGIN

Query Match 80.8%; Score 19.4; DB 10; Length 183349;
 Best Local Similarity 95.2%; Pred. No. 93; Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	2	AGCCACCAAGGAGCCAGA	22	VCCNA/c	VCCNA	1603 bp	mRNA	linear	PLN 03-OCT-1994
Db	97066	AGCCACCAAGGAGCCAGA	97086	LOCUS	V. cardunculus mRNA for cynarase.				
RESULT 14									
AL80918		AL05918		DEFINITION	X. cardunculus mRNA for cynarase.				
LOCUS				ACCESSION	X69193				
DEFINITION				VERSION	X69193.1				
ACCESSION				KEYWORDS	GT:499015				
VERSION				SOURCE					
JOURNAL				ORGANISM					
COMMENT				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
				JOURNAL					
				REFERENCE					
				AUTHORS					
				TITLE					
	</td								

مکتبہ
سیدنا
حاج شاہ
دین احمد

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 20:19:46 ; Search time 70 Seconds
(without alignments)

190.269 Million cell updates/sec

Title: US-09-831-272-7

Perfect score: 24
Sequence: 1 cagccaccaaaaggagccagaat 24

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27747546 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cggn2_6/ptodata/a/2/ina/5A_COMB.seq: *
2: /cggn2_6/ptodata/a/2/ina/5B_COMB.seq: *
3: /cggn2_6/ptodata/a/2/ina/6A_COMB.seq: *
4: /cggn2_6/ptodata/a/2/ina/6B_COMB.seq: *
5: /cggn2_6/ptodata/a/2/ina/pctus_COMB.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c 1	19	79.2	3692 2	US-09-5895-987-17
c 2	19	79.2	3692 2	US-09-5895-976-17
c 3	17.4	72.5	2606 2	US-08-5809-976-17
c 4	16.6	69.2	289 4	US-09-433-184-111
c 5	16.6	69.2	349 4	US-09-833-381-1287
c 6	16.6	69.2	2543 4	US-09-976-547-575
c 7	16.6	69.2	2636 4	US-09-976-547-383
c 8	16.2	67.5	375 4	US-09-702-705-961
c 9	16.2	67.5	375 4	US-09-702-705-1158
c 10	16.2	67.5	375 4	US-09-736-57-961
c 11	16.2	67.5	375 4	US-09-736-457-1158
c 12	16.2	67.5	375 4	US-09-614-24B-961
c 13	16.2	67.5	375 4	US-09-614-2AB-1158
c 14	16.2	67.5	375 4	US-09-671-325-961
c 15	16.2	67.5	571 4	US-09-702-705-921
c 16	16.2	67.5	571 4	US-09-736-457-921
c 17	16.2	67.5	571 4	US-09-614-124B-921
c 18	16.2	67.5	571 4	US-09-671-325-921
c 19	16.2	67.5	571 4	US-09-671-325-921
c 20	16.2	67.5	791 4	US-09-702-705-1805
c 21	16.2	67.5	791 4	US-09-736-457-1805
c 22	16.2	67.5	791 4	US-09-671-325-1805
c 23	16.2	67.5	1833 1	US-08-554-720A-36
c 24	16.2	67.5	3188 3	US-08-943-731-183
c 25	16.2	67.5	20084 3	US-09-943-731-5
c 26	16.2	67.5	889 4	US-09-227-357-88
c 27	16.7	66.7	1665 4	US-08-812-008-31

ALIGNMENTS

RESULT 1	US-08-588-983-17/c	Sequence 54, Appl
	Sequence 17, Application US/08588983	Sequence 6, Appl
	Patent No. 5854067	Sequence 8, Appl
	GENERAL INFORMATION:	Sequence 10, Appl
	APPLICANT: Christopher B. Newgard, et al.	Sequence 11, Appl
	TITLE OF INVENTION: Methods and Compositions	Sequence 12, Appl
	NUMBER OF SEQUENCES: 43	Sequence 13, Appl
	CORRESPONDENCE ADDRESS:	Sequence 14, Appl
	ADDRESSEE: Arnold, White & Durkee	Sequence 15, Appl
	STREET: P.O. Box 4433	Sequence 16, Appl
	CITY: Houston	Sequence 17, Appl
	STATE: TX	Sequence 18, Appl
	ZIP: 77210	Sequence 19, Appl
	COUNTRY: US	Sequence 20, Appl
	COMPUTER READABLE FORM:	Sequence 21, Appl
	MEDIUM TYPE: Floppy disk	Sequence 22, Appl
	OPERATING SYSTEM: PC-DOS/MS-DOS	Sequence 23, Appl
	SOFTWARE: PatentIn Release #1.0, Version #1.30	Sequence 24, Appl
	CURRENT APPLICATION DATA:	Sequence 25, Appl
	APPLICATION NUMBER: US/08/588,983	Sequence 26, Appl
	FILING DATE: Concurrently herewith	Sequence 27, Appl
	CLASSIFICATION: 424	Sequence 28, Appl
	ATTORNEY/AGENT INFORMATION:	Sequence 29, Appl
	NAME: Fussey, Shelley P.M.	Sequence 30, Appl
	REGISTRATION NUMBER: 39,458	Sequence 31, Appl
	REFERENCE/DOCKET NUMBER: UTSD:424/FUS	Sequence 32, Appl
	TELECOMMUNICATION INFORMATION:	Sequence 33, Appl
	TELEPHONE: (512) 418-3000	Sequence 34, Appl
	TELEFAX: (512) 474-7577	Sequence 35, Appl
	TELEX: n/a	Sequence 36, Appl
	INFORMATION FOR SEQ ID NO: 17:	Sequence 37, Appl
	SEQUENCE CHARACTERISTICS:	Sequence 38, Appl
	LENGTH: 3692 base pairs	Sequence 39, Appl
	TYPE: nucleic acid	Sequence 40, Appl
	STRANDEDNESS: Single	Sequence 41, Appl
	TOPOLOGY: linear	Sequence 42, Appl
	US-08-588-933-17	Sequence 43, Appl
	Query Match Similarity 79.2%; Score 19; DB 2; Length 3692;	Sequence 44, Appl
	Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;	Sequence 45, Appl
	Matches 19; Conservative 0; Sequence 1805, Ap	Sequence 46, Appl
Qy	1 CAGCCACCAAAAGGAGGCC 19	Sequence 47, Appl
Db	1087 CAGGCCACCAAAAGGAGGCC 1069	Sequence 48, Appl

RESULT 2
US-08-588-976-17/c
Sequence 17, Application US/08588976
; Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christoper B. Newgard et al
TITLE OF INVENTION: Methods and Compositions for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTS/481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 3692 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: linear
US-08-588-976-17

Query Match 79.2%; Score 19; DB 2; Length 3692;
Best Local Similarity 100.0%; Pred. No. 12; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatches 0; Indels 0; Gaps 0;

Qy	Db	Length	Score	DB	Similarity	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps
1	1087	3692	19	2	79.2%	100.0%	12	0	0	0

RESULT 3
US-08-701-233B-1/c
Sequence 1, Application US/08701233B
; Patent No. 5861308
GENERAL INFORMATION:
APPLICANT: Freudchuh, Michael; Renner, Christoph
TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS ASSOCIATED WITH T CELL ACTIVATION AND USES THEREOF
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADRESSE: Felfe & Lynch
STREET: 805 Third Avenue
CITY: New York City
STATE: New York
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
COMPUTER: IBM
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/701,233B
FILING DATE: 21-August-1995

RESULT 4
US-09-441-184-111/c
Sequence 111, Application US/09443184A
; Patent No. 6372431
GENERAL INFORMATION:
APPLICANT: Cunningham, Mary Jane
APPLICANT: Zweiger, Gary
APPLICANT: Kaser, Matthew R.
APPLICANT: Panzer, Scott
APPLICANT: Seilhamer, Jeffrey J.
APPLICANT: Yue, Henry
APPLICANT: Baughn, Mariah
APPLICANT: Azimai, Yalda
APPLICANT: Lal, Preeti
TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS
FILE REFERENCE: PC-0007 US
CURRENT APPLICATION NUMBER: US/09/443,184A
CURRENT FILING DATE: 1999-11-19
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL program
SEQ ID NO: 111
LENGTH: 289
TYPE: DNA
FEATURE: misc-feature
ORGANISM: Rattus norvegicus
FEATURE: misc-feature
NAME/KEY: misc-feature
OTHER INFORMATION: Incyte ID No. 6372431 700067537H1
US-09-443-184-111

Query Match 69.2%; Score 16.6; DB 4; Length 289;
Best Local Similarity 82.6%; Pred. No. 1.1e+02; Mismatches 4; Indels 0; Gaps 0;
Matches 19; Conservative 0; MisMatches 4; Indels 0; Gaps 0;

Qy	Db	Length	Score	DB	Similarity	Best Local Similarity	Pred. No.	Mismatches	Indels	Gaps
1	267	289	23	4	69.2%	82.6%	1.1e+02	4	0	0

RESULT 5
US-09-833-381-1287/c
Sequence 1287, Application US/09833391
; Patent No. 6672186
GENERAL INFORMATION:
APPLICANT: Robison, Keith E.
TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
FILE REFERENCE: 5/00-119
CURRENT APPLICATION NUMBER: US/09/833,381
CURRENT FILING DATE: 2001-04-11

```

; PRIORITY APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 1287
; LENGTH: 349
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc_feature
; NAME/KEY: misc_feature
; LOCATION: (1..(349))
; OTHER INFORMATION: n = A,T,C or G
; US-09-833-381-1287

; Query Match 69.2%; Score 16.6; DB 4; Length 349;
; Best Local Similarity 82.6%; Pred. No. 1.1e+02; Indels 0; Gaps 0;
; Matches 19; Conservative 0; Mismatches 4;
Qy 1 CAGCCACCAAGAGGACCGAGAA 23
Db 45 CAGCACACCAGAGGCCAGTA 23

; RESULT 6
; US-09-976-594-575/C
; Sequence 575, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furtress, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976, 594
; PRIORITY FILING DATE: 2000-10-12
; PRIORITY APPLICATION NUMBER: 60/240,409
; PRIORITY FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 575
; LENGTH: 2543
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 1859458CB1
; US-09-976-594-575

; Query Match 69.2%; Score 16.6; DB 4; Length 2543;
; Best Local Similarity 82.6%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
; Matches 19; Conservative 0; Mismatches 4;
Qy 1 CAGCCACCAAGAGGACCGAGAA 23
Db 2081 CAGACACCAAGAGGCCAGTA 23

; RESULT 7
; US-09-976-594-383
; Sequence 383, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furtress, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976, 594
; PRIORITY FILING DATE: 2000-10-12
; PRIORITY APPLICATION NUMBER: 60/240,409
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; LENGTH: 2636

; Query Match 69.2%; Score 16.6; DB 4; Length 2636;
; Best Local Similarity 82.6%; Pred. No. 1.4e+02; Indels 0; Gaps 0;
; Matches 19; Conservative 0; Mismatches 4;
Qy 1 CAGCCACCAAGAGGACCGAGAA 23
Db 906 CAGCAGSCCAAGAGGCCAGTA 928

; RESULT 8
; US-09-702-705-961/C
; Sequence 961, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Ranger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TREATMENT OF CANCER
; FILE REFERENCE: 210121-478C14
; CURRENT APPLICATION NUMBER: US/09/702, 705
; CURRENT FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FASTSEQ For Windows Version 3.0
; SEQ ID NO 961
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-702-705-961

; Query Match 67.5%; Score 16.2; DB 4; Length 375;
; Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
; Matches 18; Conservative 0; Mismatches 3;
Qy 2 AGCCACCAAGAGGACCGAGAA 22
Db 163 AGCCACCAAGAGGACCAA 143

; RESULT 9
; US-09-702-705-1158
; Sequence 1158, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Ranger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TREATMENT OF CANCER
; FILE REFERENCE: 210121-478C14
; CURRENT APPLICATION NUMBER: US/09/702, 705
; CURRENT FILING DATE: 2000-03-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSEQ for Windows Version 3.0
; LENGTH: 2636

```

```

; SEQ ID NO: 1158
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-1158

Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 3; Organism: Homo sapien
Qy 2 AGCCACCAAGAGGAGCCAGA 22
Db 213 AGCCACCAAGAGGAGCCAGA 233

RESULT 10
US-09-736-457-961/c
; Sequence 961, Application US/0976457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736, 457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 961
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-961

Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 3; Organism: Homo sapien
Qy 2 AGCCACCAAGAGGAGCCAGA 22
Db 163 AGCCACCAAGAGGAGCCAA 143

RESULT 11
US-09-736-457-1158
; Sequence 1158, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736, 457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SEQ ID NO: 1158
; LENGTH: 375
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 1158
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-1158

Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 3; Organism: Homo sapien
Qy 2 AGCCACCAAGAGGAGCCAGA 22
Db 163 AGCCACCAAGAGGAGCCAA 143

RESULT 12
US-09-614-124B-961/c
; Sequence 961, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/514, 24B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 961
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-961

Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 3; Organism: Homo sapien
Qy 2 AGCCACCAAGAGGAGCCAGA 22
Db 163 AGCCACCAAGAGGAGCCAA 143

RESULT 13
US-09-614-124B-1158
; Sequence 1158, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614, 24B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 1158
; LENGTH: 375
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO: 1158
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-1158

Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 18; Conservative 0; Mismatches 3; Organism: Homo sapien
Qy 2 AGCCACCAAGAGGAGCCAGA 22
Db 213 AGCCACCAAGAGGAGCCAA 233

```

```

; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-614-124B-1158
Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Mismatches 3;
Matches 18; Conservative 0; Indels 0; Gaps 0;
Qy 2 AGCCACCAAGGAGCCAGA 22
Db 213 AGCCACCAAGGAGCCAA 233

```

```

RESULT 14
US-09-671-325-961/c
; Sequence 961, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Rettler, Marc
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 961
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-671-325-961

```

```

; ORGANISM: Homo sapien
; US-09-671-325-1158
Query Match 67.5%; Score 16.2; DB 4; Length 375;
Best Local Similarity 85.7%; Pred. No. 1.7e+02; Mismatches 3;
Matches 18; Conservative 0; Indels 0; Gaps 0;
Qy 2 AGCCACCAAGGAGCCAGA 22
Db 213 AGCCACCAAGGAGCCAA 233

```

Search completed: July 30, 2004, 22:06:01
Job time : 72 secs

```

RESULT 15
US-09-671-325-1158
; Sequence 1158, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Rettler, Marc
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1158
; LENGTH: 375
; TYPE: DNA
; ORGANISM: Homo sapien
; US-09-671-325-961

```

(Order) *Yester* *Open*, *Open*

Copyright (c) 1993 - 2004 Compugen Ltd.

GenCore version 5.1.6

Om nucleic - nucleic search, using sw model

Run on: July 30, 2004, 20:26:46 ; Search time 406 Seconds
(without alignments)

289.841 Million cell updates/sec

Title: US-09-831-272-7

Perfect score: 24

Sequence: 1 cagccaccaaaaggagccagaat 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 245157024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cggn2_6/ptodata/2/pubpna/us07_PUBCOMB.seq:*

2: /cggn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cggn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cggn2_6/ptodata/2/pubpna/us06_PUBCOMB.seq:*

5: /cggn2_6/ptodata/2/pubpna/us07_NEW_PUB.seq:*

6: /cggn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cggn2_6/ptodata/2/pubpna/us08_NEW_PUB.seq:*

8: /cggn2_6/ptodata/2/pubpna/us09_PUBCOMB.seq:*

9: /cggn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*

10: /cggn2_6/ptodata/2/pubpna/us09C_PUBCOMB.seq:*

11: /cggn2_6/ptodata/2/pubpna/us09C_PUBCOMB.seq:*

12: /cggn2_6/ptodata/2/pubpna/us09_NEW_PUB.seq:*

13: /cggn2_6/ptodata/2/pubpna/us09c_PUBCOMB.seq:*

14: /cggn2_6/ptodata/2/pubpna/us10A_PUBCOMB.seq:*

15: /cggn2_6/ptodata/2/pubpna/us10C_PUBCOMB.seq:*

16: /cggn2_6/ptodata/2/pubpna/us10c_PUBCOMB.seq:*

17: /cggn2_6/ptodata/2/pubpna/us10c_PUBCOMB.seq:*

18: /cggn2_6/ptodata/2/pubpna/us10c_PUBCOMB.seq:*

19: /cggn2_6/ptodata/2/pubpna/us10c_PUBCOMB.seq:*

20: 15 19 79.2 18772 10 US-09-764-861-63 Sequence 63, Appl

c 16 19 79.2 18772 13 US-09-764-861-63 Sequence 63, Appl

c 17 19 79.2 18772 15 US-10-15-928-63 Sequence 63, Appl

c 18 18.2 75.8 625 13 US-10-027-632-211873 Sequence 211873,

c 19 18.2 75.8 625 13 US-10-027-632-211873 Sequence 211874,

c 20 18.2 75.8 625 16 US-10-027-632-211873 Sequence 211873,

c 21 18.2 75.8 625 16 US-10-027-632-211874 Sequence 211874,

c 22 18.2 75.8 2689 16 US-10-260-238-74 Sequence 474, App

c 23 18.2 75.8 3515 17 US-10-437-963-91409 Sequence 91409, A

c 24 18.2 75.8 36211 16 US-10-085-117-49 Sequence 49, Appl

c 25 17.8 74.2 41 12 US-10-035-833A-123 Sequence 123, App

c 26 17.8 74.2 41 12 US-10-035-833A-5462 Sequence 5462, Appl

c 27 17.8 74.2 436 13 US-10-236-392-85 Sequence 83, Appl

c 28 17.8 74.2 446 13 US-10-236-392-83 Sequence 89, Appl

c 29 17.8 74.2 529 13 US-10-236-392-89 Sequence 271, Ap

c 30 17.6 73.3 255 12 US-09-732-627A-2711 Sequence 92107, A

c 31 17.6 73.3 567 13 US-10-027-632-92107 Sequence 92108, A

c 32 17.6 73.3 567 16 US-10-027-632-92107 Sequence 92107, A

c 33 17.6 73.3 567 16 US-10-027-632-92108 Sequence 92108, A

c 34 17.6 73.3 567 16 US-10-027-632-92108 Sequence 92108, A

c 35 17.6 73.3 624 13 US-10-07-632-187229 Sequence 187229,

c 36 17.6 73.3 624 16 US-10-027-632-187229 Sequence 187229,

c 37 17.6 73.3 693 13 US-10-282-1122A-19603 Sequence 19603, A

c 38 17.6 73.3 746 10 US-09-764-891-1197 Sequence 197, Ap

c 39 17.6 73.3 79467 12 US-10-052-882-223 Sequence 223, App

c 40 17.6 73.3 162025 13 US-09-834-700-13 Sequence 13, Appl

c 41 17.6 73.3 162025 13 US-09-834-700-14 Sequence 14, Appl

c 42 17.6 73.3 162025 13 US-09-834-700-17 Sequence 17, Appl

c 43 17.6 73.3 162025 13 US-09-834-700-18 Sequence 18, Appl

c 44 17.6 73.3 162025 15 US-10-272-665-35 Sequence 35, Appl

c 45 17.6 73.3 162025 15 US-10-272-665-36 Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-09-770-423-146/c

; Sequence 146, Application US/09770423

; Publication No. US2002004490A1

; GENERAL INFORMATION:

; APPLICANT: Gorlach, Jorn

; APPLICANT: An, Yong-Qiang

; APPLICANT: Hamilton, Carol M.

; APPLICANT: Price, Jennifer L.

; APPLICANT: Raine, Tracy M.

; APPLICANT: Yu, Yang

; APPLICANT: Rameka, Joshua G.

; APPLICANT: Page, Amy

; APPLICANT: Abraham V.

; APPLICANT: Ledford, Brook L.

; APPLICANT: Woessner, Jeffrey P.

; APPLICANT: Haas, William David

; APPLICANT: Garcia, Carlos A.

; APPLICANT: Kricker, Maja

; APPLICANT: Slader, Ted

; APPLICANT: Davis, Keith R.

; APPLICANT: Hoffman, Neil

; APPLICANT: Hurlan, Patrick

; TITLE OF INVENTION: Expressed Sequences of Arabidopsis thaliana

; TITLE OF INVENTION: thaliana

; FILE REFERENCE: 2028 (PARA-017PRV)

; CURRENT APPLICATION NUMBER: US/09-770-423

; CURRENT FILING DATE: 2001-01-26

; PRIOR APPLICATION NUMBER: 60/178,512

; PRIOR FILING DATE: 2000-01-27

; NUMBER OF SEQ ID NOS: 999

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 146 LENGTH: 431

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-770-423-146

Query Match 82.5%; Score 19.8; DB 13; Length 431;
 Best Local Similarity 91.3%; Pred. No. 38; Mismatches 2;
 Matches 21; Conservative 0; Indels 0; Gaps 0;

RESULT 2
US-10-424-599-65132

Sequence 65132, Application US/01424599
 Publication No. US2004031072A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa Thomas J
 APPLICANT: Kovacic David K
 APPLICANT: Zhou Yihua
 APPLICANT: Cao Yongwei
 TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With Title of Invention: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223)B

CURRENT APPLICATION NUMBER: US/10/424,599

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO: 65132

LENGTH: 1451

TYPE: DNA

ORGANISM: Glycine max

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_29827C.1
US-10-424-599-65132

Query Match 82.5%; Score 19.8; DB 13; Length 1451;
 Best Local Similarity 91.3%; Pred. No. 37; Mismatches 21; Conservative 0; Indels 0; Gaps 0;

Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 AGCCACCAAGGAGGCCAGAT 24
Db 367 AACCAACAAAGGAGCCAAAT 345RESULT 3
US-09-938-842A-496/C

Sequence 496, Application US/09938842A
 Patent No. US20020160378A1
 GENERAL INFORMATION:

APPLICANT: Harper, Jeff
 APPLICANT: Kreps, Joel

APPLICANT: Zhu, Tong

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPT300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO: 496

LENGTH: 2868

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842A-496

RESULT 4
US-09-938-842B-496/C

Sequence 495, Application US/09938842A
 Publication No. US20040009476A9
 GENERAL INFORMATION:
 APPLICANT: Harper, Jeff
 APPLICANT: Kreps, Joel

APPLICANT: Wang, Xun

TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

TITLE OF INVENTION: SAME, AND METHODS OF USE

FILE REFERENCE: SCRIPT1300-3

CURRENT APPLICATION NUMBER: US/09/938,842A

CURRENT FILING DATE: 2001-08-24

PRIOR APPLICATION NUMBER: US 60/227,866

PRIOR FILING DATE: 2000-08-24

PRIOR APPLICATION NUMBER: US 60/264,647

PRIOR FILING DATE: 2001-01-16

PRIOR APPLICATION NUMBER: US 60/300,111

PRIOR FILING DATE: 2001-06-22

NUMBER OF SEQ ID NOS: 5379

SEQ ID NO: 496

LENGTH: 2868

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-09-938-842B-496

RESULT 5
US-10-350-923B-45/C

Sequence 45, Application US/10350923B
 Publication No. US20040002084A1
 GENERAL INFORMATION:
 APPLICANT: Dotson, Stanton B.

APPLICANT: Ma, Xiao Jun

TITLE OF INVENTION: Nucleic acid, polypeptides, vectors, and cells derived from activ

FILE REFERENCE: S03157-01

CURRENT APPLICATION NUMBER: US/10/350,923B

CURRENT FILING DATE: 2003-01-24

PRIOR APPLICATION NUMBER: US/09/454,280

PRIOR FILING DATE: 1999-12-03

PRIOR APPLICATION NUMBER: PCT/US99/28773

PRIOR FILING DATE: 1999-06-12

PRIOR APPLICATION NUMBER: US 60/111,006

PRIOR FILING DATE: 1998-12-04

NUMBER OF SEQ ID NOS: 79

SOFTWARE: Patentin version 3.1

SEQ ID NO: 45

LENGTH: 141

TYPE: DNA

ORGANISM: Homo sapiens

US-10-350-923B-45

Query Match 79.2%; Score 19.1; DB 16; Length 141;
 Best Local Similarity 100.0%; Pred. No. 89; Mismatches 0; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCACCAAAGGAGCC 19

```

Db      102 |||||||CAGCCACCAAGGAGGACCC 84
; Sequence 3382, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIORITY APPLICATION NUMBER: US 60/211,379
; PRIORITY FILING DATE: 2000-06-14
; PRIORITY APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3382
; LENGTH: 3062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 US1333
; US-09-880-107-3382

Query Match          79.2%; Score 19; DB 9; Length 3062;
Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; ; mismatches 0; indels 0; gaps 0;

Qy      1 CAGCCACCAAGGAGGACCC 19
Db      1081 CAGCCACCAAGGAGGACCC 1063

RESULT 7
US-10-342-887-728/c
; Sequence 728, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van de Velde, Laura
; APPLICANT: Van de Velver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIORITY APPLICATION NUMBER: 60/380,770
; PRIORITY FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 728
; LENGTH: 3062
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESION NUMBER: NM 002115
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-728

Query Match          79.2%; Score 19; DB 13; Length 3062;
Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; ; mismatches 0; indels 0; gaps 0;

Qy      1 CAGCCACCAAGGAGGACCC 19
Db      1081 CAGCCACCAAGGAGGACCC 1063

RESULT 8
US-10-172-118-728/c
; Sequence 728, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Velde, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIORITY APPLICATION NUMBER: 60/380,770
; PRIORITY FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 728
; LENGTH: 3062
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESION NUMBER: NM 002115
; DATABASE ENTRY DATE: 2001-06-18
; US-10-172-118-728

Query Match          79.2%; Score 19; DB 13; Length 3062;
Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; ; mismatches 0; indels 0; gaps 0;

Qy      1 CAGCCACCAAGGAGGACCC 19
Db      1081 CAGCCACCAAGGAGGACCC 1063

RESULT 9
US-10-717-397-214/c
; Sequence 214, Application US/10717597
; Publication No. US20040110221A1
; GENERAL INFORMATION:
; APPLICANT: Wreth
; APPLICANT: Burczynski, Michael E.
; APPLICANT: Twine, Natalie C.
; APPLICANT: Dorner, Andrew J.
; APPLICANT: Trepicchio, William L.
; APPLICANT: Slonim, Donna K.
; APPLICANT: Stover, Jennifer A.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING RCC AND OTHER SOLID TUMORS
; FILE REFERENCE: AM01080L
; CURRENT APPLICATION NUMBER: US/10/717,597
; CURRENT FILING DATE: 2003-11-21
; PRIORITY APPLICATION NUMBER: US 60/459,782
; PRIORITY FILING DATE: 2003-04-03
; PRIORITY APPLICATION NUMBER: US 60/427,982
; PRIORITY FILING DATE: 2002-11-21
; NUMBER OF SEQ ID NOS: 4904
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 214
; LENGTH: 3062
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESION NUMBER: NM 002115
; DATABASE ENTRY DATE: 2003-04-03
; US-10-717-597-214

Query Match          79.2%; Score 19; DB 13; Length 3062;
Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; ; mismatches 0; indels 0; gaps 0;

```

Best Local Similarity 100.0%; Pred. No. 82; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCACCAAGAGGACCC 19
Db 1081 CAGCCACCAAGAGGACCC 1063

RESULT 10
US-10-177-809-1/c
; Sequence 1, Application US/10177809
; Publication No. US20030148305A1
GENERAL INFORMATION:
APPLICANT: Gross, William M.
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Burgess, Catherine E.
APPLICANT: Bader, Joel S.
APPLICANT: Bansal, Aruna
APPLICANT: Pena, Carol E.A.
APPLICANT: Shimkets, Richard A.
APPLICANT: Kekuda, Ramesh
APPLICANT: Zerhusen, Bryan
APPLICANT: Smithson, Glenna
APPLICANT: Anderson, David
APPLICANT: Zhong, Mei
APPLICANT: Miller, Charles
APPLICANT: Vernet, Corine
APPLICANT: Hjalt, Tord
TITLE OF INVENTION: NUCLEIC ACIDS, POLYPEPTIDES, SINGLE NUCLEOTIDE POLYMORPHISMS AND
TITLE OF INVENTION: OF USE THEREOF
FILE REFERENCE: 21402-390
CURRENT APPLICATION NUMBER: US/10/177,809
CURRENT FILING DATE: 2003-01-28
PRIORITY APPLICATION NUMBER: 60/311,285
PRIORITY FILING DATE: 2001-08-09
PRIORITY APPLICATION NUMBER: 60/299,949
PRIORITY FILING DATE: 2001-06-21
PRIORITY APPLICATION NUMBER: 60/300,290
PRIORITY FILING DATE: 2001-06-22
PRIORITY APPLICATION NUMBER: 60/302,883
PRIORITY FILING DATE: 2001-07-03
PRIORITY APPLICATION NUMBER: 60/327,345
PRIORITY FILING DATE: 2001-10-05
PRIORITY APPLICATION NUMBER: 60/327,467
PRIORITY FILING DATE: 2001-10-05
PRIORITY APPLICATION NUMBER: 60/327,892
PRIORITY FILING DATE: 2001-10-09
PRIORITY APPLICATION NUMBER: 60/354,586
PRIORITY FILING DATE: 2002-02-05
NUMBER OF SEQ ID NOS: 16
SOFTWARE: Custom
SEQ ID NO 1
LENGTH: 3205
TYPE: DNA
ORGANISM: Homo sapiens
US-10-177-809-3
Query Match 79.2%; Score 19; DB 15; Length 3205;
Best Local Similarity 100.0%; Pred. No. 82; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCACCAAGAGGACCC 19
Db 1081 CAGCCACCAAGAGGACCC 1063

RESULT 12
US-10-177-809-5/c
; Sequence 5, Application US/10177809
; Publication No. US20030148305A1
GENERAL INFORMATION:
APPLICANT: Gross, William M.
APPLICANT: Alsobrook, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Burgess, Catherine E.
APPLICANT: Bader, Joel S.
APPLICANT: Bansal, Aruna
APPLICANT: Pena, Carol E.A.
APPLICANT: Shimkets, Richard A.
APPLICANT: Kekuda, Ramesh
APPLICANT: Zerhusen, Bryan
APPLICANT: Smithson, Glenna
APPLICANT: Anderson, David

RESULT 11
US-10-177-809-3/c
; Sequence 3, Application US/10177809

; APPLICANT: Zhong, Mei
 ; APPLICANT: Miller, Charles
 ; APPLICANT: Vernet, Corinne
 ; APPLICANT: Hjalt, Tord
 ; TITLE OF INVENTION: NUCLEIC ACIDS, POLYPEPTIDES, SINGLE NUCLEOTIDE POLYMORPHISMS AND FILE REFERENCE: 21402-390
 ; CURRENT APPLICATION NUMBER: US/10/177,809
 ; PRIOR APPLICATION NUMBER: 60/311,285
 ; PRIOR FILING DATE: 2001-08-09
 ; PRIOR APPLICATION NUMBER: 60/299,949
 ; PRIOR FILING DATE: 2001-06-21
 ; PRIOR APPLICATION NUMBER: 60/300,290
 ; PRIOR FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: 60/302,883
 ; PRIOR FILING DATE: 2001-07-03
 ; PRIOR APPLICATION NUMBER: 60/327,345
 ; PRIOR FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 60/327,467
 ; PRIOR FILING DATE: 2001-10-05
 ; PRIOR APPLICATION NUMBER: 60/327,892
 ; PRIOR FILING DATE: 2001-10-09
 ; PRIOR APPLICATION NUMBER: 60/354,586
 ; PRIOR FILING DATE: 2002-02-05
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: Custom
 ; SEQ ID NO: 5
 ; LENGTH: 3205
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-177-809-5

 Query Match 79.2%; Score 19; DB 15; Length 3205;
 Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CAGCCACCAAGAGGACCC 19
 Db 1081 CAGCCACCAAGAGGACCC 1063

 RESULT 13
 US-10-152-319A-1767/C
 ; Sequence 1767, Application US/10152319A
 ; Publication No. US20040072160A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mendrick, Donna
 ; APPLICANT: Porter, Mark
 ; APPLICANT: Johnson, Kory
 ; APPLICANT: Higgs, Brandon
 ; APPLICANT: Castle, Arthur
 ; APPLICANT: Elashoff, Michael
 ; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
 ; FILE REFERENCE: 43921-5090US
 ; CURRENT APPLICATION NUMBER: US/10/191,803
 ; PRIOR FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,819
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/305,623
 ; PRIOR FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: US 60/369,351
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US 60/377,611
 ; PRIOR FILING DATE: 2002-05-06
 ; NUMBER OF SEQ ID NOS: 1140
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 138
 ; LENGTH: 3692
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE: OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_022179
 ; US-10-191-803-138

 Query Match 79.2%; Score 19; DB 16; Length 3692;
 Best Local Similarity 100.0%; Pred. No. 82; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 CAGCCACCAAGAGGACCC 19
 Db 1087 CAGCCACCAAGAGGACCC 1069

 RESULT 14
 US-10-191-803-138/C
 ; Sequence 138, Application US/10191803
 ; Publication No. US20040014040A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MENDRICK, Donna
 ; APPLICANT: PORTER, Mark
 ; APPLICANT: JOHNSON, Kory
 ; APPLICANT: HIGGS, Brandon
 ; APPLICANT: CASTLE, Arthur
 ; APPLICANT: ELASHOFF, Michael
 ; TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
 ; FILE REFERENCE: 43921-5090US
 ; CURRENT APPLICATION NUMBER: US/10/191,803
 ; PRIOR FILING DATE: 2002-07-10
 ; PRIOR APPLICATION NUMBER: US 60/303,819
 ; PRIOR FILING DATE: 2001-07-10
 ; PRIOR APPLICATION NUMBER: US 60/305,623
 ; PRIOR FILING DATE: 2001-07-17
 ; PRIOR APPLICATION NUMBER: US 60/369,351
 ; PRIOR FILING DATE: 2002-04-03
 ; PRIOR APPLICATION NUMBER: US 60/377,611
 ; PRIOR FILING DATE: 2002-05-06
 ; NUMBER OF SEQ ID NOS: 1140
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 138
 ; LENGTH: 3692
 ; TYPE: DNA
 ; ORGANISM: Rattus norvegicus
 ; FEATURE: OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_022179
 ; US-09-764-861-63/C
 ; Sequence 63, Application US/09764861
 ; Publication No. US20030171252A9
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; PRIOR APPLICATION NUMBER: US 60/324,928
 ; PRIOR FILING DATE: 2001-09-27

```
; FILE REFERENCE: PT204
; CURRENT APPLICATION NUMBER: US/09/764,861
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 63
; LENGTH: 1872
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9782)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-861-63
```

```
Query Match 79.2%; Score 19; DB 10; Length 18772;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 CAGCCACCAAGAGGACCC 19
          ||||| ||||| ||||| |
Db      708 CAGCCACCAAGAGGACCC 690
```

Search completed: July 30, 2004, 22:13:04
Job time : 410 secs

GenCore version 5.1.6
 Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 30, 2004, 20:13:41 ; Search time 2963 Seconds

(without alignments)
 241.881 Million cell updates/sec

Title: US-09-831-272-7

Perfect score: 24

Sequence: 1 cagccaccaaaggaggaccagaat 24

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of bits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_estbum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estol:*

7: em_estro:*

8: em_htc:*

9: gb_esti:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_main:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

ALIGNMENTS

RESULT 1

CK269862/c

LOCUS CK269862

DEFINITION EST715940 potato abiotic stress cDNA library

clone PolCP48 5' end, mRNA sequence.

ACCESSION CK269862

VERSION CK269862.1

SOURCE EST

ORGANISM Solanum tuberosum (potato)

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamiales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 934)

AUTHORS Buell,C.R., Hart,A., Zismann,V., Karamcheyeva,S.A. and Baker,B.

TITLE Generation of ESTs from abiotic stressed potato tissue

JOURNAL Unpublished (2003)

COMMENT Contact: Robin Buell

The Institute for Genomic Research
 9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from TIGR via potato@tigr.org

Seq Primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers

/organism="Solanum tuberosum"

/mol_type="mRNA"

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
c 1	20.4	85.0	934 14 CK269862	CK269862 EST715940
c 2	19.8	82.5	347 29 CC468543	CC468543 CH240 139
c 3	19.8	82.5	528 9 AV52210	AV52210 AV52210
c 4	19.8	82.5	531 28 AZ114521	AZ114521 RPCI-23-4

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

/cultivar="Kennebec"
 /db_xref="taxon:4113"

/clones="POACP8"
 /tissue_type="Abiotic stress treated leaf and root tissue"

/lab_host="DHB0B-Tona"
 /clone_lib="Biotron abiotic stress cDNA library"

/notes="Vector: PCMVSPORT6.1; Site_1: BCOI; Site_2: NOLI;
 supplier: Solanum tuberosum var. Kennebec plants were

grown from cuttings on a 16hr light/8 hr dark cycle at 25
 C for 3-4 weeks. Abiotic stress conditions were applied to

four separate sets of plants. Set 1 involved saturation of
 the soil with 150 mM NaCl and tissues were harvested after

following application of the salt stress (leaves: 2hr, 6hr,
 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).

Set 2 were grown under the standard conditions and then
 were water stressed by withdrawal of further watering

applications. Drought stressed plants were harvested after

cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
 and 5d). Set 3 were grown under the standard conditions
 and then were cold stressed by placement at 4 C. Cold
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d. Set 4 were grown under the standard conditions and
 then were heat stressed by placement at 35 C. Heat
 stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
 2d and 4d and heat-stressed roots were harvested at 6 hr,
 12 hr, 1 d, and 4d. RNA was isolated from all tissues and
 equal RNA from each tissue and stress was pooled to
 construct the cDNA library. RNA sample."

ORIGIN

Query Match 85.0%; Score 20.4; DB 14; Length 934;
 Best Local Similarity 95.5%; Pred. No. 4.8e+02;
 Matches 21; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AGCCACCAAGAGGAGCCAGAA 23
 Db 171 AGGCCACCAAGAGGAGCCAGAA 150

RESULT 2
 CC68543/c
 LOCUS CC68543 347 bp DNA linear GSS 12-JUN-2003
 DEFINITION CH240_139N06 TV CHORI-240 Bos taurus genomic clone CH240_139N06,
 genomic survey sequence.
 ACCESSION CC68543
 VERSION CC68543.1 GI:31654775
 KEYWORDS GSS.

SOURCE Bos taurus (cow)
 ORGANISM Bos taurus

REFERENCE
 AUTHORS Costa,J.N., Mota,M. and Caetano,A.R.
 TITLE Brazil's Contribution to End-Sequencing the Bovine BAC Library
 JOURNAL Unpublished (2003)
 COMMENT Contact: Caetano AR
 Department of Biotechnology
 Embrapa Recursos Genéticos e Biotecnologia
 Parque Estação Biológica, Final Av. W/5 NORTE, Brasília-DF C.P.
 02772-0770-9000 Brasil
 Tel.: 55 61 448 4778
 Fax: 55 61 340 3658
 Email: acastano@cenargen.embrapa.br

Clones are derived from the bovine BAC library CHORI-240 (<http://www.chori.org/bacpac/bacpac.html>). Bases shown have Phred quality value equal to or higher than 20. Bases with quality value below 20 were masked with 'N'. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources

(http://www.chori.org/bacpac/ordering_information.htm). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBC) by Embrapa Recursos Genéticos e Biotecnologia with financing from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil.

Plate: 139 row: N column: 06
 Seq primer: T7
 Class: BAC ends

Location/Qualifiers
 1..347

FEATURES Source
 High quality sequence stop: 347.

Query Match 82.5%; Score 19.8; DB 29; Length 347;
 Best Local Similarity 91.3%; Pred. No. 6.4e+02;
 Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 AGCCACCAAGAGGAGCCAGAA 24
 Db 298 AGCCACCAAGAGGAGCCAGAA 276

RESULT 3
 AV529210/c
 LOCUS AV529210
 DEFINITION AV529210 Arabidopsis thaliana aboveground organs two to six week old Arabidopsis thaliana CDNA clone APZL30d07R 5', mRNA sequence.
 ACCESSION AV529210
 VERSION AV529210.1 GI:8689493
 KEYWORDS EST
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Bokarvata; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS
 TITLE
 A large scale analysis of cDNA in *Arabidopsis thaliana*: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
 JOURNAL DNA Res. 7 (3), 175-180 (2000)
 MEDLINE 20363093
 PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The FIRST Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL:<http://www.kazusa.or.jp/en/plant/>
 Location/Qualifiers

FEATURES source
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="APZL30d07R"
 /tissue_type="aboveground organ"
 /dev_stage="two to six week old"
 /clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
 /note="Vector: pBlueScriptII SK-, site_1: BCOI; site_2: XbaI"

ORIGIN

QY	2	AGCCACCAAGAGGAGCCAGAT	24	RESULT 4	AZ114521	AZ114521	531 bp	DNA	linear	GSS 12-MAY-2000	Best Local Similarity 91.3%; Score 82.5%;	Score 19.8; DB 9;	length 528;	EST 01-SEP-2000
Db	44	AACCACCAAGAGGAGCCAAAT	22	DEFINITION	RPCI-23-460E1.TV	RPCI-23	Mus musculus	genomic clone	RPCI-23-460E1,	LOCUS	AV529714/C	594 bp	mRNA	linear EST 01-SEP-2000
REFERENCE				ACCESSION	AZ114521	genomic survey sequence.				DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				VERSION	AZ114521.1	GI:7774836				VERSION	AV529714	Old Arabidopsis thaliana cDNA clone AP2L44g05R 5'	,	mRNA sequence.
TITLE				JOURNAL						KEYWORDS	AV529714.1	GI:8889997		EST
JOURNAL				SOURCE						ORGANISM	Arabidopsis thaliana	(thale cress)		
COMMENT				ORGANISM						Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicots; core eudicots; rosids; euroids II; Brassicales; Brassicaceae; Arabidopsis.				
ORGANISM				KEYWORDS						REFERENCE	1	(bases 1 to 534)		
Mus musculus		(house mouse)		KEYWORDS						AUTHORS	Asamizu,E., Nakamura,Y., Satoh,S. and Tabata,S.			
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				KEYWORDS						TITLE	A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries.			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				KEYWORDS						JOURNAL	Kazusa DNA Research Institute			
1 (bases 1 to 531)				KEYWORDS						MEDLINE	Kazusa DNA Research Institute			
Zhao,S., Nieman,W., Pfeilblyum,T., Malek,J., Shatsman,S., Akinret,B., Levins,M., Megann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Fraser,C.M.				KEYWORDS						PUBMED	Kazusa DNA Research Institute			
Mouse BAC End Sequences from Library RPCI-23				KEYWORDS						COMMENT	10907847	Contact: Erika Asamizu		
Unpublished (1999)				KEYWORDS						COMMENT	20363093	The First Laboratory for Plant Gene Research		
Other_GSS: RPCI-23-460E1.TJ				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Contact: Shuying Zhao				KEYWORDS						COMMENT	10907847	Kazusa DNA Research Institute		
Department of Bukaryotic Genomics				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
The Institute for Genomic Research				KEYWORDS						COMMENT	10907847	Kazusa DNA Research Institute		
9712 Medical Center Dr., Rockville, MD 20850, USA				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Fax: 301 838 0208				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Email: szhao@tigr.org				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@ddj.org med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderframe.html) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tgb/bac_ends/mouse/bac_end_intro.html				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Plate: 460	row: E	column: 1		KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Seq Primer: T7				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Class: BAC ends.				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
FEATURES				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
source				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
1. -531				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Location/Qualifiers				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
1. /organism="Mus musculus"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/mol_type="genomic DNA"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/strain="C57BL/6J"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/db_xref="taxon:00090"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/clone="RPCI-23-460E1"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/sex="Female"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/lab_host="BRL10B"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
/clone_lib="RPCI-23"				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Note="Organ: Kidney; Brain; Vector: pBACE3.6; Site_1: EcorI; Site_2: EcoRI; Site_3: EcoRI C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the DR10B sites. The ligation products were transformed into electrocompetent cells (BRL Life Technologies). ORIGIN				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
Query Match 82.5%; Score 19.8; DB 28; Length 531; Best Local Similarity 91.3%; Pred. No. 7.3e+02; Mismatches 21; Conservative 0; Indels 0; Gaps 0;				KEYWORDS						COMMENT	20363093	Kazusa DNA Research Institute		
QY	1	CAGCCACCAAGAGGAGCGAA	23	RESULT 6	CK317567/c	CK317567	705 bp	mRNA	linear	EST 16-DEC-2003	Best Local Similarity 91.3%; Score 82.5%;	Score 19.8; DB 9;	length 534;	EST 16-DEC-2003
Db	150	CAGCCACCAAGAGGAGCGAA	172	DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714/C	594 bp	mRNA	linear EST 01-SEP-2000
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Old Arabidopsis thaliana cDNA clone AP2L44g05R 5'	,	mRNA sequence.
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
AUTHORS				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
TITLE				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
JOURNAL				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
COMMENT				DEFINITION	VRJ376	Vitis riparia	endoendormant bud	- VRJ	Vitis riparia	DEFINITION	AV529714	Arabidopsis thaliana	aboveground organs	two to six-week
REFERENCE				DEFINITION	VRJ376</									

		FAX: 605 688 4713 Email: Anne_Fennell@sdsstate.edu Seq primer: PolyT Location/Qualifiers
FEATURES	source	/organism="Vitis riparia" 1. .705 /mol_type="mRNA" /db_xref="taxon:96939" /clone="VRJ376" /dev_stage="dormant" /lab_host="DH10B" /clone lib="Vitis riparia endodormant bud - VRJ7" /note="Organ: bud; Vector: pSport 1; Site_1: SAlI; Site_2: NotI; VRJ is a cDNA library of Vitis riparia endodormant buds. Endodormant buds were collected from insect and disease free vines that were induced into dormancy with short photoperiods. The directionally oriented library was constructed according to Gibco BRL Superscript Plasmid System for cDNA synthesis and plasmid cloning."
		ORIGIN
		Query Match 82.5%; Score 19.8; DB 14; Length 705; Best Local Similarity 91.3%; Pred. No. 9e+02; Mismatches 2; Matches 21; Conservative 0; Mismatches 2; Qy 2 AGCCACCAAGGAGGCCAGAT 24 Db 297 AGCCACGAAGGAGGCCAGCT 275
RESULT 7	BG024130	LOCUS BE729686 mRNA linear EST 15-SEP-2000 DEFINITION 60156283F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3832425 5', DEFINITION mRNA sequence. DEFINITION NIH-MGC http://mgc.ncbi.nih.gov/ DEFINITION National Institutes of Health, Mammalian Gene Collection (MGC) DEFINITION Unpublished (1999) DEFINITION Contact: Robert Strauberg, Ph.D. DEFINITION Email: cgabbs@mail.nih.gov DEFINITION Tissue Procurement: ATCC/DCTD/DTP DEFINITION cDNA Library Preparation: Ling Hong/Rubin Laboratory DEFINITION CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DEFINITION DNA Sequencing by: Incyte Genomics, Inc. DEFINITION Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov DEFINITION Plate: LLM509 row: j column: 10 DEFINITION High Quality sequence stop: 737.
FEATURES	source	REFERENCE AUTHORS TITLE JOURNAL COMMENT
VERSION	BG024130.1	REFERENCE 1 (bases 1 to 1082) AUTHORS NIH-MGC http://mgc.ncbi.nih.gov/ TITLE NIH-MGC http://mgc.ncbi.nih.gov/ JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC) COMMENT Unpublished (1999) Contact: Robert Strauberg, Ph.D. Email: cgabbs@mail.nih.gov Tissue Procurement: ATCC/DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov Plate: LLM509 row: j column: 10 High Quality sequence stop: 737.
KEYWORDS		
ORGANISM	Homo sapiens (human)	
REFERENCE		
AUTHORS		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE		
JOURNAL		National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT		Unpublished (1999)
COMMENT		Contact: Robert Strauberg, Ph.D. Email: cgabbs@mail.nih.gov
TISSUE		Tissue Procurement: ATCC
PROCUREMENT		CDNA Library Preparation: Life Technologies, Inc.
LIBRARY		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
PREPARATION		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
LIBRARY		CDNA Library Prepared by: Life Technologies, Inc.
ARRAYED		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BY		CDNA Library Prepared by: Life Technologies, Inc.
INCYTE GENOMICS		CDNA Library Prepared by: Incyte Genomics, Inc.
INCYTE GENOMICS		CDNA Library Prepared by: Incyte Genomics, Inc.
CONTACT		Contact: Robert Strauberg, Ph.D.
EMAIL		Email: cgabbs@mail.nih.gov
TISSUE		Tissue Procurement: ATCC
PROCUREMENT		CDNA Library Preparation: Life Technologies, Inc.
LIBRARY		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
ARRAYED		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BY		CDNA Library Prepared by: Incyte Genomics, Inc.
INCYTE GENOMICS		CDNA Library Prepared by: Incyte Genomics, Inc.
CONTACT		Contact: Robert Strauberg, Ph.D.
EMAIL		Email: cgabbs@mail.nih.gov
TISSUE PROCUREMENT		Tissue Procurement: ATCC
LIBRARY PREPARATION		CDNA Library Preparation: Life Technologies, Inc.
LIBRARY ARRAYED		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
BY INCYTE GENOMICS		CDNA Library Prepared by: Incyte Genomics, Inc.
CONTACT		Contact: Robert Strauberg, Ph.D.
EMAIL		Email: cgabbs@mail.nih.gov
TISSUE DISTRIBUTION		Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
CONTACT		http://image.llnl.gov
EMAIL		Plate: LLM1090 Row: e column: 02
HIGH QUALITY SEQUENCE STOP		High quality sequence stop: 610.
LOCATION/QUALIFIERS		Location/Qualifiers
FEATURES	source	ORIGIN
		Query Match 82.5%; Score 19.8; DB 10; Length 1082; Best Local Similarity 91.3%; Pred. No. 9e+02; Mismatches 2; Matches 21; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
		Qy 2 AGCCACCAAGGAGGCCAGAT 24 Db 834 AGCCACGAAGGAGGCCGAAT 856
RESULT 9	AV21379	LOCUS AV271379 RIKEN full-length enriched, adult male testis (DH10B) Mus DEFINITION AV271379 RIKEN full-length enriched, adult male testis (DH10B) Mus DEFINITION musculus cDNA clone 4931408J02 3', mRNA sequence. DEFINITION AV271379 DEFINITION AV271379.1 GI:6259416 DEFINITION EST. DEFINITION SOURCE ORGANISM Mus musculus DEFINITION SOURCE ORGANISM Mus musculus (house mouse)
FEATURES	source	ORIGIN

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 242)

REFERENCE
AUTHORS

Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Carninci,P., Endo,T.,
Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N., Hirozane,T., Hori,F.,
Ishii,Y., Ishikawa,T., Itoh,M., Izawa,M., Kadota,K., Kagawa,I.,
Kai,C., Kawai,J., Kikuchi,N., Kojima,Y., Koya,S., Kusakabe,M.,
Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,
Owa,C., Ozawa,Y., Saito,H., Sano,M., Sato,K., Shibusawa,K.,
Shibata,Y., Shigemoto,Y., Shiraki,T., Sogabe,Y., Sugihara,Y.,
Suzuki,H., Suzuki,H., Takahashi,F., Tatemoto,M., Tominga,N.,
Tsunoda,T., Watanabe,S., Yamamura,T., Yasunishi,A.,
Yokota,T., Yoshiki,A., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.

TITLE
JOURNAL

RIKEN Mouse ESTB (Kono,H., et al. 1999)

Unpublished

Comment:

Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Fax: 81-45-503-9215

Email: genome-res@gsc.riken.go.jp

URL: http://genome.gsc.riken.go.jp/

Sasaki,N., Izawa,M., Watahiki,M., Ozawa,K., Tanaka,T., Yoneda,Y.,
Matsuraya,S., Carninci,P., Muramatsu,M., Okazaki,Y. and
Hayashizaki,Y.

Transcriptional sequencing: A method for DNA sequencing using RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. 95 (7), 3455-3460 (1998)

Itoh,M., Kitsunai,T., Akiyama,J., Shibata,K., Izawa,M., Kawai,J.,
Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M.,
Okazaki,Y. and Hayashizaki,Y.

Automated filtration-based high-throughput plasmid preparation

System. Genome Res. 9 (5), 463-470 (1999)

Carninci,P. and Hayashizaki,Y.

High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
19-44 (1999)

Please visit our web site (<http://genome rtc.riken.go.jp>) for

further details.

location/Qualifiers

source

1. .242

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="Taxon:10090"

/clone="493140802"

/sex="male"

/tissue_type="testis"

/dev_stage="adult"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, adult male testis

/note="Site_1: Sali; Site_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN. Division of Experimental Animal Research in Riken

contributed to prepare mouse tissues. 1st strand cDNA was

primed with a primer [5'

GAGAGAGAGAGAGATCCAGAGCTGTTTGTGTTTGN 3'], cDNA was

prepared by using trehalose thermo-activated reverse

transcriptase and subsequently enriched for full-length by

cap-trapper. Second strand cDNA was prepared with the

primer adapter of sequence [5'

GAGAGAGAGATCCAGAGCTGTTTGTGTTTGN 3']. cDNA

was cloned into the XbaI and BamHI sites. Vector: a

modified pBluescript KS(+) after bulk excision from Lambda

FLC 1. Cloning sites, 5' end: Sali; 3' end: BamHI."

ORIGIN

Query Match

80.0%

Score 19.2;

DB 9;

Length 242;

best local similarity 87.5%;

Pred. No. 1e+03;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

Conservative 0;

Mismatches 3;

Indels 0;

Gaps 0;

Matches 21;

<p

contributed to prepare mouse tissues. 1st strand cDNA was printed with a primer [5'-GAGGAGGAGGATCCAGAGCTTGTGTTTGTGTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5' - GAGGAGGAGTCTCGAGTATTAAATTATCCTCCCCCCCCC 3']. cDNA was cloned into the XbaI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda EIC I. Cloning sites, 5', end: SalI; 3', end: BamHI"

ORIGIN

Query Match 80.0%; Score 19.2; DB 9; Length 257; Best Local Similarity 87.5%; Pred. No. 1.e+03; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0; Qy 1 CAGGCCACCAAGAGGACCCGAAT 24 Db 54 CAGCCCATCACAGGACCCAGTT 77

ORIGIN

Query Match 80.0%; Score 19.2; DB 13; Length 708; Best Local Similarity 87.5%; Pred. No. 1.4e+03; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0; Qy 1 CAGGCCACCAAGAGGACCCGAAT 24 Db 147 CAGSCCAAAGAGGAAACCGAAT 170

ORIGIN

RESULT 11
BU026771 BU026771 708 bp mRNA linear EST 23-AUG-2002
DEFINITION OG17P22.yg abl OH ERGHJ sunflower RHA280 Helianthus annuus cDNA clone OG17P22, mRNA sequence.
ACCESSION BU026771
VERSION BU026771.1
KEYWORDS EST.
ORGANISM Helianthus annuus (common sunflower)
Bulkyota; Vridiplante; Streptophyta; Embryophyta; Tracheophyta; Asparidales; campanulids; Asterales; Asteraceae; Asteridae; Heliantheae; Helianthus.
1. (bases 1 to 708)

REFERENCE Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L., Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison,P., Kolkman,J., Slabaugh,M.S., Livingston,Y., Lai,Z., Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
JOURNAL Unpublished (2002)
COMMENT Department of Vegetable Crops, R.W.Michelmore]
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530) 742-1742
Fax: 1-(530) 752-8659
Email: akozik@ucdavis.edu [michelmore@veggmail.ucdavis.edu]
Singleton, see <http://cgpdbs.ucdavis.edu/> for details.
Plate: OG17 row: P column: 22.

FEATURES source
1. loci
/organism="Helianthus annuus"
/mol_type="mRNA"
/cultivar="RHA280"
/db_xref="taxon:4232"
/clone="OG17P22"
/lab_host="E. coli"
/clone_lib="OH_ERGHJ sunflower RHA280"
/note="Vector: pBRCDNASIB: The library was constructed from 11 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgpdbs.ucdavis.edu/>

TAG_TISSUE=chemical induction
TAG_LIB=OH_ERGHJ_sunflower_RHA280
TAG_SEQ=TGTAGCCGGG"

ORIGIN

Query Match 80.0%; Score 19.2; DB 13; Length 708; Best Local Similarity 87.5%; Pred. No. 1.4e+03; Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0; Qy 1 CAGGCCACCAAGAGGACCCGAAT 24 Db 147 CAGSCCAAAGAGGAAACCGAAT 170

ORIGIN

RESULT 12
AZ312990/C AZ312990 728 bp DNA linear GSS 29-SEP-2000
LOCUS IM0029H08F Mouse 10kb plasmid UGCGM library Mus musculus genomic clone UGCGM0029H08 F, genomic survey sequence.
REFERENCE A2312990
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beaumont,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenah,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D. Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: dunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0029 Row: H column: 08
Seq primer: CGTGCTAACGACGGCACT
Class: plasmid ends
High quality sequence stop: 728.
Location/Qualifiers

1. . 728
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGCGM0029H08"
/sex="Male"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGCGM library"
/note="Vector: pMD2N; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD2 (gi|47321149|AF29072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to

adapted vector DNA, and transformed into chemically-competent *E. coli* XLI-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match Best Local Similarity 80.0%; Score 19.2; DB 28; Length 728;
Matches 21; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CAGCCACCAAGAGGACCCAGAT 24
Db 58 CAGACACCAAATAGCACCAAGAT 35

RESULT 13

DEFINITION CC518089 CH240_364J21_T7 CHORI-240 Bos taurus genomic clone CH240_364J21,
ACCESION CC518089
VERSION CC518089.1 GI:31836377
KEYWORDS GSS.
SOURCE Bos taurus (cow)

ORGANISM

Bos taurus; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovidae; Bovinae; Bos.

REFERENCE

1 (bases 1 to 899)
Holt, R., Stott, J., Yang, G., Barber, S., Smailius, D., Prabhu, A.-L.,

Tsai, M., Cloutier, J., Lee, D., Gurn, N., Olson, T., Mayo, M.,

Butterfield, Y., Kirkpatrick, R., Liu, J., Guin, R., Chan, A., Chiu, R.,

Matthewson, C., Wye, N., Mason, A., Brown-John, M., Jones, S.,

Schein, J., Marr, M., de Jong, P., McWilliam, S., Barries, W.,

Dalrymple, B.P. and Tellam, R.

Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398

Unpublished (2003)
COMMENT JOURNAL

Contact: Rob Holt
Other GSAs: CH240_364J21.TARBC13P2

The British Columbia Cancer Agency Genome Science Centre

600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6

Tel: 604-877-6085
Fax: 604-877-6276

Email: rholte@bcgsc.ca

Copies are derived from the bovine BAC library CHORI-240

(<http://www.chori.org/bacpac/bovine240.html>). For BAC library

availability, please contact Pieter de Jong (pdejong@mail.cho.org).

Copies may be purchased from BACPAC Resources

(<http://www.chori.org/bacpac/ordering-information.html>). This work

was undertaken as part of the International Bovine BAC Mapping

Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the

British Columbia Genome Sciences Centre, Canada.

Plate: 364 Row: J Column: 21

Seq primer: T7

Class: BAC ends

Location/Qualifiers

1. .899

/organism="Bos taurus"

/mol_type="genomic DNA"

/strain="breed: Hereford"

/db_xref="Taxon:9913"

/clone="CH240_364J21"

/sex="Male"

/cell_type="Blood"

/clone_libr="CHORI-240"

/note="Vector: pBAC1.3; Site_1: SmaI; Site_2: MboI; Hereford bull Li Domino 93375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

FEATURES

source

RESULT 14
DEFINITION BF901013/c
LOCUS PM1-MT0203-081200-003-b03 MT0203 Homo sapiens cDNA, mRNA sequence.
VERSION BF901013.1 GI:12292472
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 359)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Steinlein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
COMMENT shotgun sequencing of the human transcriptome with ORP expressed sequence tags
PROC. NATL. ACAD. SCI. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL 2002663
MEDLINE 10737800
PUBMED Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM1&t2=PM1-MT0203-081200-003-b03&t3=2001-12-08&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 359.
Location/Qualifiers

FEATURES
source

1. .359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stge="Adult"
/clone_libr="Mbo203"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) Profiles from the PUC 18 vector. Reverse transcription of tissue into the PUC 18 vector. mRNA and cDNA amplification were performed under low stringency conditions."

RESULT 15
DEFINITION CD557470/c
LOCUS AGENCOURT_14413595 NIH_MGC_180 Homo sapiens cDNA clone

Query Match Best Local Similarity 79.2%; Score 19; DB 10; Length 359;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAGCCACCAAGAGGACCC 19
Db 137 CAGCCACCAAGAGGACCC 119

IMAGE:3038677 5', mRNA sequence.
 ACCESSION CD557470
 VERSION CD557470.1 GI:31583538
 KEYWORDS EST:
 SOURCE Homo sapiens (human)
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 (bases 1 to 897)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics /
 National Cancer Institute / NIH
 Bldg. 31 Rm1007 Bethesda, MD 20892
 Email: cgapbsr@mail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 cDNA Library Preparation: Invitrogen Corp
 cDNA Library Arrayed: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing By: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: NDAM458 Row: f Column: 06
 High quality sequence stop: 627.
 FEATURES source
 Location/Qualifiers 1..897
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ttxon:9606"
 /clone="IMAGE:3038677"
 /lab_host="DHIOB-Ton A ("T1 and T5 phage resistances)"
 /clone_libs="NIH_MGC_180"
 /note="Organ: Testis; Vector: pCMW-SPORT6.1; Site 1: NotI;
 Site 2: EcoRV (destroyed); Library is oligo-dT primed and
 directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.68 kb. Library was
 constructed by (Invitrogen). Note: this is a NIH_MGC
 library."
 ORIGIN
 Query Match 79.2%; Score 19; DB 14; Length 897;
 Best Local Similarity 100.0%; Pred. No. 1.8e+03; Mismatches 0;
 Matches 19; Conservative 0; Gaps 0;
 Indels 0;
 QY 1 CAGGCCACCAAGAGGACCC 119
 DB 417 CAGGCCACCAAGAGGACCC 399
 Search completed: July 30, 2004, 22:04:43
 Job time : 2973 SECs